

Aus dem Institut für medizinisch-biologische
Statistik der Universität Marburg/Lahn
Direktor: Prof. Dr. rer. nat. P. Ihm

Analyse einer Batterie von Tests,
die zur Aufnahme von Studienfachwechslern
in die Medizinische Fakultät der Philipps-Universität Marburg/Lahn
durchgeführt wurden

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der gesamten Medizin der Philipps-Universität Marburg
Fachbereich Humanmedizin

vorgelegt von

Ilse Strempel
aus Paderborn

Marburg 1973

Druck- und Verlagsanstalt

Werner Blasaditsch · 89 Augsburg · Theodor-Heuß-Platz 3 · Telefon (0821) 25563

Printed in Germany

Aus dem Institut für medizinisch-biologische
Statistik der Universität Marburg/Lahn
Direktor: Prof. Dr. rer. nat. P. Ihm

Analyse einer Batterie von Tests, die zur Aufnahme von
=====
Studienfachwechslern in die Medizinische Fakultät der
=====
Philipps-Universität Marburg/Lahn durchgeführt wurden
=====

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der gesamten Medizin der Philipps-Universität Marburg
Fachbereich Humanmedizin

vorgelegt von

Ilse Strempel
aus Paderborn

Marburg 1973



Angenommen vom Fachbereich Humanmedizin der
Universität Marburg am 21.8.1973

gedruckt mit Genehmigung des Fachbereichs

Direktor: Prof.Dr.med.H.Oepen (Dekan)

Referent: Prof.Dr.rer.nat.P.Ihm

Correferent: Prof.Dr.med.Dr.rer.nat.H.Remschmidt

Meinen Eltern in
Dankbarkeit gewidmet



Digitized by the Internet Archive
in 2026

Inhaltsverzeichnis

	Seite
I. Einleitung	1
II. Fragestellung	2
III. Methodik	3
1. Kollektiv	3
2. Angewandte Leistungskriterien	3
a) Schulnoten	3
b) Klausuren	4
α) Botanik	6
β) Zoologie	6
γ) Chemie	7
δ) Physik	7
c) Intelligenz-Struktur-Test (I-S-T) von AMTHAUER	8
d) Vorphysikumsnoten	10
IV. Diskussion der Methodik	11
1. Stichprobe	11
2. Schulnoten und Marburger Punkte (MP)	11
3. Naturwissenschaftliche Klausuren	13
a) Multiple-Choice-Tests	13
b) Tests mit freier Beantwortung	15
4. Diskussion des I-S-T-AMTHAUER	16
5. Vorphysikumsnoten	17
V. Statistische Auswertungsmethoden	18
1. Mittelwertberechnung	18
2. Normalverteilung	18
3. Korrelationskoeffizienten	18
Besondere Test- und Itemstatistiken	
a) Die Validität	21
b) Die Reliabilität	22
c) Der Trennschärfeindex (TI)	23
d) Der Schwierigkeitsindex (SI)	23

	Seite
4.	Die Faktorenanalyse 24
5.	Hauptkomponentenanalyse 26
VI.	Darstellung der Ergebnisse 27
1.	Klausurstatistiken 27
2.	Rohwerteverteilungen 28
a)	Botanik 28
b)	Zoologie 29
c)	Physik 30
d)	Chemie 31
3.	Vergleich der SI (Häufigkeitsverteilungen) 32
a)	Botanik 32
b)	Zoologie 32
c)	Physik und Chemie 33
4.	Vergleich der Trennschärfeindizes 34
5.	Interkorrelationsmuster der Klausuren 35
a)	Botanik 35
b)	Zoologie 36
c)	Chemie 37
d)	Physik 38
6.	Interne Hauptkomponentenanalysen der Klausuren 39
a)	Botanik 39
b)	Zoologie 40
c)	Physik 41
d)	Chemie 42
7.	Korrelationen zwischen Klausurergebnissen und Abiturnoten 43
8.	Interkorrelationen der Intelligenztestergebnisse und Korrelationen der I-S-T-AMTHAUER-Items mit den Leistungen in den 4 Klausuren 44
9.	Mittelwerte und Streuungen der Leistungen der zum Vorphysikum gelangten Humanmediziner 45
10.	Mittelwerte und Streuungen der Abiturnoten aller 140 Bewerber 46
11.	Interkorrelationsmuster der Abiturzeugnisnoten 46

	Seite
12.	Ergebnisse des I-S-T-AMTHAUER 47
	MW und Streuungen der Items 47
a)	aller Bewerber 47
b)	über- und unterdurchschnittlicher Pbn getrennt nach Klausurleistungen 48
c)	über- und unterdurchschnittlicher Pbn getrennt nach Vorphysikumsleistungen 49
	Darstellung der Intelligenzprofile 50
d)	aller Bewerber im Vergleich zum Ärzteprofil AMTHAUERS 50
e)	über- und unterdurchschnittlicher Pbn getrennt nach Klausurleistungen 51
f)	über- und unterdurchschnittlicher Pbn getrennt nach Vorphysikumsleistungen 52
g)	Darstellung der internen HKA des I-S-T-AMTHAUER 53
13.	Korrelationen der 4 Leistungskomplexe 54
a)	als Gesamtkomplexe 54
b)	mit ihren einzelnen Leistungsvariablen 55
14.	Hauptkomponentenanalysen 56
a)	mit 6 Variablen 56
b)	mit 13 Variablen 59
c)	mit 22 Variablen 57
VII.	Darstellung der Ergebnisse 60
1.	Leistungskomplex Klausuren 60
a)	Botanik 60
b)	Zoologie 61
c)	Physik 61
d)	Chemie 62
2.	Leistungskomplex Abitur 63
3.	Leistungskomplex I-S-T-AMTHAUER 64
4.	Leistungskomplex: Vorphysikum 65
5.	Statistische Analyse der Wechselbeziehungen des gesamten Variablenkollektivs 65

	Seite
VIII.	Diskussion der Ergebnisse 69
1.	Leistungskomplex Klausuren 69
a)	Botanikklausur 69
b)	Zoologieklausur 76
c)	Physikklausur 78
d)	Chemieklausur 81
e)	Wechselbeziehungen der Klausuren mit externen Variablen 82
α)	Klausuren und I-S-T-AMTHAUER 82
β)	Beziehungen Klausuren/Abitur 83
γ)	Beziehungen Klausuren/Vorphysikum 84
2.	Leistungskomplex Abitur 85
a)	Beziehungen der Noten zum Vorphysikum 85
b)	Beziehungen der Abiturnoten zueinander und Vergleiche mit der Literatur 86
c)	Beziehungen zwischen Schulleistungen und Intelligenz 87
3.	Leistungskomplex: Vorphysikum 88
4.	Leistungskomplex I-S-T-AMTHAUER 89
a)	Beziehungen zwischen Intelligenz und den übrigen Leistungskomplexen 89
b)	Diskussion der Intelligenzprofile 89
c)	Interne HKA des I-S-T 90
5.	Faktorenstruktur des Variablenkollektivs 92
IX.	Zusammenfassung 95
X.	Literaturverzeichnis 97
XI.	Anhang 103
	Originaltext der Botanikklausur 103
	Statistiken der Botanikklausur 107
	Originaltext der Zoologieklausur 113
	Itemstatistik der Zoologieklausur 118
	"Wechsler-Klausur" vom 29. Juni 1968 (Physik) 124
	Statistiken der Physikklausur 128
	Originaltext der Chemieklausur 129
	Statistiken der Chemieklausur 131

I. Einleitung:

Die Bedeutung der medizinisch-didaktischen Forschung ist erst in den letzten Jahren mit der Intensivierung der Diskussionen um neue Ausbildungsformen in vollem Umfang erkannt worden. Die Bemühungen um eine effektivere Ausbildung der Studenten, erfordern eine völlige Neuordnung und Straffung des Medizinstudiums. Das wurde notwendig durch:

- a) die ständig wachsende Fülle der Wissensinhalte
- b) die "steigende Verkürzung der Halbwertszeit unseres Wissens" (vgl. UEXKUELL 1968)
- c) die Notwendigkeit einer verstärkten Ausrichtung der Ausbildung auf die praktischen Anforderungen im Beruf
- d) die stark ansteigenden Studentenzahlen.

Die dadurch notwendig gewordenen differenzierteren Prüfungsmethoden wurden bereichert durch Einbeziehung psychologischer Testmethoden. Es boten sich hierfür besonders schriftliche Leistungskontrollen mit Multiple-Choice-Fragestellungen an, für die im amerikanischen Schrifttum schon langjährige Erfahrungen vorliegen (HUBBARD und CLEMANS 1960; GARDNER 1968; CHAUNCEY 1968; LINDQUIST 1951; DAVE 1968; TITTLE 1968; KAPUSTE und NOACK 1968).

Derartige Tests werden in allen staatlichen Prüfungen obligat sein, nach Möglichkeit jedoch auch im gesamten Ausbildungsverlauf zur Überprüfung des Wissensstandes am Ende bestimmter Ausbildungsphasen angewendet werden. Die Probleme der Leistungsmessung und die Bedeutung von Leistungstests sind in verschiedenen Arbeiten bereits behandelt worden. Ausführliche Ergebnisse liegen besonders im Bereich der Schule vor (INGENKAMP und MARSOLEK 1968; WEINGARD 1964). Besondere Probleme wurden in Deutschland in Folge der Zulassungsbeschränkung für verschiedene Studiengänge hervorgerufen. Dieses gilt insbesondere für das Studium der Medizin. Seine spezielle Zulassungssproblematik betreffend, liegen eingehende Untersuchungen vor über den Wert der Reifeprüfung und ihrer Zeugnisnoten als

Prädiktor für späteren Studienerfolg (KAPUSTE 1966; SCHÜTZ 1966; TESCHNER 1965; INGENKAMP und MARSOLEK 1967).

Die neuen Formen der Hochschulprüfungen befinden sich in Deutschland jedoch überwiegend noch im Stadium des Experimentes (HÖHN 1968):

"Es wird heute sehr viel über Prüfungen positiv wie negativ ausgesagt, was niemals exakt untersucht worden ist und daher bestenfalls den Charakter einer Hypothese hat. Hier ist empirische Forschung, die die Zuverlässigkeit und Bewährung verschiedener Auslesemethoden kontrolliert und zuverlässigere Meßinstrumente schafft, dringend nötig."

Die hier vorliegende Arbeit will versuchen, diesbezüglich einen Beitrag zu liefern.

II. Fragestellung

Die spezielle Themenstellung dieser Arbeit ergab sich aus der günstigen Situation, daß von einem größeren Studentenkollektiv die Ergebnisse verschiedener, zeitlich aufeinander folgender Leistungskontrollen aus Schule und Hochschule, sowie eine psychologische Intelligenzmessung vorlagen. Bei diesem Datenmaterial war von besonderem Interesse:

1. welche qualitativen Aussagen über die angewandten Hochschulleistungstests anhand spezieller Testuntersuchungsverfahren gemacht werden können
2. welche Beziehungen besonders der Hochschulleistungen zu den übrigen Variablen bestehen
3. welche Faktorenstruktur die Faktorenanalyse für das gesamte Variablenkollektiv ermittelt.

III. Methodik

1. Kollektiv:

Das untersuchte Datenmaterial wurde am Ende des SS 1968 an der Philipps-Universität in Marburg an 140 Studenten der naturwissenschaftlichen Fakultät ermittelt. Von diesen Bewerbern entfielen 87 auf die Humanmedizin und 53 auf die Zahnmedizin. Sie kamen aus 4 verschiedenen Schultypen, nämlich zu 50 % von neu-sprachlichen, zu 15 % von altsprachlichen, zu 30 % von naturwissenschaftlichen Gymnasien, und 6 % der Bewerber hatten das Abitur über den zweiten Bildungsweg erreicht. Dieselben Studenten hatten sich zuvor vergeblich um Aufnahme zum Studium der Medizin oder Zahnmedizin beworben. Ihnen wurden nach 1- bis 2-semesterigem naturwissenschaftlichem Studium eine erneute Zulassungsmöglichkeit geboten durch Teilnahme an der hier untersuchten Testbatterie.

2. Angewandte Leistungskriterien:

Die Analysen gelten für die 4 großen Leistungsgruppen: Schulnoten, Klausuren, I-S-T-AMTHAUER und Vorphysikum.

a) Schulnoten:

In Marburg war das entscheidende Zulassungskriterium zur Aufnahme des Medizinstudiums, die sogenannte "Marburger-Punktezahl" (MP). Es handelt sich dabei um eine gewichtete Summe aus den Noten in vier Kernfächern des Abiturs und zwar Deutsch, Mathematik, Physik und Chemie. Der für jeden Bewerber errechnete Punktwert wurde dabei nach folgendem Verfahren ermittelt:

Notenskalen im Reifezeugnis Punkte für die einzelnen Noten

1	1,0
2	1,6
3	2,2
4	2,8
5	3,4
6	4,0

Summe der Punkte	Punktzahl
4,0	20
4,5	19
5,0	18
5,5	17
6,0	16
7,0	14
7,5	13
8,0	12
8,5	11
9,0	10
9,5	9
10,0	8
10,5	6
11,0	4
11,5	2
12,0	0

Die MP sind als Einzelvariable in den unten durchgeführten Untersuchungen vertreten. Außerdem wurden die Zensuren folgender Schulfächer als Einzelvariable in die Analysen mit einbezogen:

- | | | |
|------------|---------------|-----------------------|
| 1. Deutsch | 2. Mathematik | 3. Gemeinschaftskunde |
| 4. Physik | 5. Chemie. | |

Eine kleine gesonderte Korrelationsmatrix aller 140 Studienbewerber erfaßt zusätzlich zu obigen folgende Fächer:

- | | | |
|----------|-------------|--------------|
| 6. Sport | 7. Englisch | 8. Biologie. |
|----------|-------------|--------------|

b) Klausuren:

Über die Zulassung der sich erneut bewerbenden Studenten wurde im wesentlichen anhand der Ergebnisse sogenannter "Wechslerklausuren" entschieden. Es handelt sich dabei um eine von der Fakultät angeordnete naturwissenschaftliche Prüfung für Bewerber zum Medizinstudium. Sie bestand aus 4 Klausuren, die nur zum SS 1968 und ausschließlich für diesen Probandenkreis konzipiert wurden. Die 4 Fächer sind gleich den vorklinischen Prüfungsfächern im Vorphysikum, und zwar Botanik, Zoologie, Physik und Chemie. Alle Prüflinge hatten

die vorausgegangenen Vorlesungen in diesen Studienfächern gehört. Die Klausuren wurden an einem Tag mit zeitlicher Limitierung geschrieben.

Die Art und Weise der Aufgabenstellung lag in der Hand der Professoren und war ihnen freigestellt. Bei den unterschiedlichen Möglichkeiten der Testkonstruktion entschieden sich die Professoren der Physik und Chemie für die bisher herkömmliche Art schriftlicher Klausuren, in denen Fragen von den Studenten frei, d.h. ohne vorgegebene Antwortmöglichkeiten, schriftlich beantwortet werden.

Die Professoren in Botanik und Zoologie dagegen wählten die modernere Art der Aufgabenkonstruktion in Form sogenannter Multiple-Choice bzw. Mehrfachwahl-Aufgaben und bildeten Multiple-Choice-Tests mit 20 Aufgaben.

Bei Multiple-Choice-Tests handelt es sich um ein Verfahren, das sich aus der Testpraxis der Psychologie herleitet, und das in den USA erstmalig in großem Umfang auch für Leistungskontrollen im Medizinstudium übernommen und weiterentwickelt wurde (HUBBARD 1968). Ein derartiger Multiple-Choice-Test besteht aus einer Anzahl von Testeinheiten, sogenannten Items, die als Mehrfachwahlaufgaben konstruiert sind. Die Antworten, meistens fünf, sind dem Prüfling vorgegeben, und er hat sich für eine der Antwortmöglichkeiten zu entscheiden, und zwar je nach Art des Aufgabentyps für die richtige, falsche oder beste Lösung oder er wird aufgefordert, bestimmte Zuordnungen zu treffen. Bei den hier vorliegenden Klausuren in Botanik und Zoologie wurden nur die drei der gewöhnlich am häufigsten angewandten Aufgabentypen ausgewählt, und zwar folgende:

a) Richtigantwort-Aufgaben,

bei denen zu einer einleitenden Frage oder einem Satz 5 Alternativen vorgegeben werden, von denen der Prüfling die richtige Antwort herauszusuchen hat (HUBBARD 1968; SCHULZ 1970).

b) Bestantwortaufgaben

hierbei hat der Proband die Alternative, die den Sachverhalt am besten ausdrückt, anzukreuzen. Gelegentlich lassen sich beide obigen Formen nicht eindeutig trennen.

c) Weilaufgaben oder sogen. "Relationship-Analysis-Items"
(HUBBARD 1968).

Sie bestehen aus 2 Sätzen, X und Y, von denen der Erste meistens eine Feststellung ist, der Zweite eine Begründung dafür gibt. Sie sind, ob logisch oder nicht, das hat der Prüfling zu entscheiden, durch "denn" oder "weil" verknüpft. Feststellung, Begründung und Verknüpfung können jeweils richtig oder falsch sein, so daß 5 Kombinationsmöglichkeiten vorhanden sind, von denen nur eine sinnvoll und richtig ist.

Es wird gefordert, daß alle angegebenen Antwortmöglichkeiten plausible Lösungen darstellen (HUBBARD 1968; SCHULZ, REMSCHMIDT, PRINZ 1970). Die von den Probanden gewählten Alternativen mußten jeweils in ein Lösungsschema eingetragen werden. Die Auswertung erfolgte per Computer.

α) Botanikklausur:

Die Botanikklausur besteht aus 20 Testitems. Jede richtige Antwort wurde mit einem Punkt bewertet. Sie war zu 45 % aus Fragen vom Typ "Richtigantwort", zu 15 % aus "Bestantwortaufgaben" und zu 40 % aus "Beziehungsaufgaben" zusammengestellt. Erfragt wurde der Inhalt des vorausgegangenen Semesterstoffes, der laut Studienplan als "allgemeine Botanik" bezeichnet wird. Es wurde dabei dem biochemischen Aspekt botanischer Probleme mit 70 % aller Fragen eine überwiegende Bedeutung beigemessen. Dabei erfragten 40 % der Textaufgaben besonders spezielle biochemische Prozesse der pflanzlichen Natur. Es bedurfte bei verschiedenen Aufgaben mehr oder weniger der Aktualisierung lexikalischen Wissens.

β) Zoologieklausur:

Auch der zweite Multiple-Choice-Test, die Zoologieklausur, besteht aus 20 Fragen, bei der jede richtige Lösung mit einem Punkt bewertet wurde. Als Richtigantwortaufgabe waren 55 % der Items formuliert, 10 % als Bestantwortaufgaben und 35 % als Beziehungsaufgaben. Inhaltlich wurde auch hier der vorausgegangene Semesterstoff erfragt. Die Hälfte der Items prüft

mehr oder weniger allgemeines zoologisches Wissen, die andere Hälfte erfordert zum Teil spezielle Kenntnisse der Vererbungslehre und Entwicklungsgeschichte. Es wird auch hier verschiedentlich lexikalisches Faktenwissen geprüft.

γ) Chemieklausur:

Der Leistungstest in Chemie bestand aus 10 in Umfang und Schwierigkeit gleichwertigen Items. Die Aufgaben erfragen zu 80 % komplizierte organisch-chemische Reaktionsabläufe. Jedes Problem erfaßt dabei ein komplexeres Geschehen, dessen Verständnis in 3 bis 8 Unteraufgaben geprüft wird. Zur vollständigen Beantwortung aller 10 Aufgaben mußten die Lösungswege von 35 zum Teil einander bedingenden Einzelproblemen dargestellt werden. Für jede vollständig richtig beantwortete Aufgabe wurden 10 Punkte vergeben, richtige Teillösungen mit entsprechend weniger Punkten bewertet. Die Auswertung wurde von den Dozenten des Institutes vorgenommen. Inhaltlich scheint die Klausur sehr homogen. Über das Abfragen reinen Faktenwissens hinaus wurde vor allem Verständnis für chemische Prozesse vorausgesetzt.

δ) Physikklausur:

Die Physikklausur mit freier Beantwortung der einzelnen Aufgaben bestand aus 10 Fragenkomplexen. Jeder einzelne entsprach dabei einem bestimmten physikalischen Problemkreis und war aus 2 bis 7 Items zusammengesetzt. Insgesamt gab es dadurch 53 Einzelaufgaben zu lösen, die je nach Schwierigkeitsgrad mit 0,5; 1,0; 1,5 oder 2,0 Punkten bewertet wurden. Die erreichbare Höchstpunktzahl war 100. Die Auswertung wurde von den Professoren und Assistenten des Institutes vorgenommen. Inhaltlich scheint diese Klausur heterogen. Es fällt auf, daß vor allem breit gestreutes Basiswissen aus verschiedenen Teilgebieten der Physik überprüft wird. Sehr spezielle physikalische Fragen wurden nicht gestellt.

Die Klausuren sind in verschiedenen Analysen sowohl als Gesamtkomplex wie auch als Einzelvariable mit ihren Gesamtpunk-

werten vertreten. Verschiedentlich wurden jedoch auch die einzelnen Subtests untersucht und in Beziehung zu anderen Variablen gesetzt.

c) Der Intelligenz-Struktur-Test von AMTHAUER:

Beim Intelligenzstrukturtest (I-S-T) geht der Autor von der Annahme aus, daß die Intelligenz eine Sonderstruktur im Gesamt der Persönlichkeitsstruktur des Menschen ist (I-S-T - AMTHAUER 1953 a, 1953 b). Diese individuelle Strukturiertheit, sowie die allgemeine Intelligenz eines Probanden oder einer Probandengruppe abschätzen und verifizieren zu können, ist Ziel dieses Testes. Ganz besonders geht es darum, die Beziehungen zwischen Intelligenz und Beruf darzulegen (PRIESTER 1964). Nach Überzeugung AMTHAUERS haben die meisten Angehörigen eines Berufes ein ganz bestimmtes feststellbares Intelligenzniveau. Für 31 Berufe ermittelte der Autor empirisch die berufstypischen Intelligenzprofile. Aus dem Vergleich des Profilverlaufs eines getesteten Probanden mit dem von AMTHAUER ermittelten Berufsprofilen können Aussagen über die Eignung dieses Bewerbers für eine bestimmte Berufsgruppe gemacht werden (AMTHAUER 1953 b). Beim I-S-T werden folgende Aspekte der Intelligenz untersucht: Sprachliche Intelligenz; rechnerische Intelligenz; räumliche Vorstellung und Reproduktion von gelernten Worten, d.h. die allgemeine Merkfähigkeit. Innerhalb der sprachlichen Intelligenz liegt der Schwerpunkt auf der Überprüfung des induktiven Denkens, des abstrakten Denkens, des kombinierenden Denkens und der Urteilsbildung (PRIESTER 1964).

Durch 9 Subtests sollen diese Seiten der Intelligenz getestet werden. Sie werden im folgenden mit den entsprechenden Eigenschaften, die sie nach AMTHAUER messen, aufgeführt:

- | | |
|-----------------------|--|
| 1. Satzergänzung (SE) | : "Urteilsbildung", "common sense", "Wirklichkeitssinn", "Selbständigkeit im Denken" |
| 2. Wortauswahl (WA) | : "Sprachgefühl", "induktives sprachliches Denken", "Erfassen von Bedeutungsgehalten", "Einfühlungsfähigkeit", |

- | | |
|-------------------------|---|
| 3. Analogien (AN) | : "Kombinationsfähigkeit",
"Erfassen von Beziehungen",
"Gründlichkeit im Denken", |
| 4. Gemeinsamkeiten (GE) | : "Abstraktionsfähigkeit",
"Begriffsbildung", "Sprachl.
Ausdrucksfähigkeit", |
| 5. Merkaufgaben (ME) | : "Merkfähigkeit", "gelernte
Wörter behalten können", |
| 6. Rechenaufgaben (RA) | : "Praktisch-rechnerisches
Denken", |
| 7. Zahlenreihen (ZR) | : "Induktives Denken mit Zahlen",
"theoret.-rechnerisches Denken", |
| 8. Figurenauswahl (FA) | : "Vorstellungsfähigkeit",
"anschaulich-ganzheitliches
Denken", |
| 9. Würfelaufgaben (WÜ) | : "Räumliches Vorstellen-Können". |

Für die einzelnen Testteile wurden standardisierte Skalenwerte berechnet. Sie gelten für das 13. - 61. Lebensjahr und für verschiedene Schulabschlüsse. Die Leistungen in den einzelnen Testeinheiten sowie das Gesamttestergebnis werden zunächst als Rohwerte ermittelt und dann mittels angegebener Tabellen in Standardwerte transformiert. Diese wiederum werden in ein Koordinatensystem (Abszisse Standardpunktwerte, Ordinate Testuntereinheiten) eingetragen. Durch Verbindung der Punkte erhält man das jeweilige Intelligenzprofil. Die Summe der Standardpunktwerte ergibt einen Anhaltspunkt für die Intelligenzhöhe eines Probanden. Sie entspricht einem ganz bestimmten IQ-Wert, einem Intelligenzmaß, das von STERN 1912 (HEISS 1964) eingeführt wurde. STERN bestimmte den IQ als Quotienten aus Intelligenzalter und Lebensalter. Während das Leistungsprofil einen Einblick in die Intelligenzstruktur ermöglicht, gibt der Gesamtstandardwert, der sich linear in IQ transformieren läßt, einen Anhalt für das Intelligenzniveau. Bei der Interpretation beider im Rahmen einer Begabungsdiagnose bedarf es jedoch großer Erfahrung, da Schulbildung, berufliche Entwicklung, soziale Herkunft usw. berücksichtigt werden müssen.

d) Vorphysikum:

Als 4. Leistungskomplex wurden die Studienerfolge der in-
zwischen in der Medizin immatrikulierten Studenten anhand
der Leistungsergebnisse ihrer ersten staatlichen Prüfung,
der naturwissenschaftlichen Vorprüfung, genannt Vorphysikum,
festgestellt. Die 4 Prüfungsfächer entsprechen dem medizinisch-
naturwissenschaftlichen vorklinischen Studium. Es handelt sich
dabei um die Fächer Physik, Chemie, Zoologie und Botanik.
Die Kenntnisse der Studenten werden jeweils in einem mündlichen
Prüfungsgespräch getestet.

Die Daten der Stichprobe von 140 Bewerbern konnten nicht in
allen Analysen verwendet werden, da die Bewerber zum Studium
der Zahnmedizin ein naturwissenschaftliches Fach weniger stu-
dieren als die Humanmediziner. Außer einer Schulnotenuntersu-
chung beziehen sich die Daten deshalb auf ein Kontingent von
84 Humanmedizinbewerbern. Da wiederum nur 36 dieser Studenten
angenommen wurden und das Vorphysikum erreichten, können Er-
folgskontrolluntersuchungen, in welche die Vorphysikumsergeb-
nisse eingehen, sich nur auf diesen Personenkreis beziehen.

Aus allen oben genannten Leistungsgruppen ergeben sich insge-
samt 84 Variable, die sowohl in Gruppen als auch einzeln unter-
sucht wurden.

IV. Diskussion der Methodik:

1) Stichprobe

Eine gewisse Homogenität des Probandenkollektivs war dadurch gewährleistet, daß die Bewerber mit ihren für Marburg bewerteten Kernfächern des Reifezeugnisses unter dem erforderlichen Zulassungsniveau lagen. Aus verschiedenen Gründen kann jedoch von einer Homogenität des Kollektivs in bezug auf die Kenntnisse und Leistungen nur bedingt ausgegangen werden. So bleibt zum Beispiel zu diskutieren, ob es signifikante regionale und schultypgebundene Unterschiede der Abiturdurchschnittszensuren gibt (KAPUSTE 1966; WEINGARDT 1964; SCHÜTZ 1968; RICHTER 1965). Darüber hinaus spielen sicherlich individuelle unterschiedliche Auffassungen von der Bedeutung der Zensuren bei den einzelnen Lehrern eine Rolle (KAPUSTE 1966; v. KÜGELGEN 1966 a, b; FURCK 1961; BERNHART 1962; KLEINERT 1951; WEINGARDT 1964).

2) Marburger Punkte (MP)

Entsprechend dem Recht der wissenschaftlichen Hochschulen "in eigener Verantwortung, ohne Einwirkung des Staates über die Zulassung zum Studium entscheiden zu können", (von STOLTENBERG 1966) erfolgt bei der Zulassung von Studienanfängern nach dem Informationsblatt über die Richtlinien der medizinischen Fakultäten der Bundesrepublik Deutschland (1966) eine Auswahl der Bewerber anhand von unterschiedlich bewerteten Notendurchschnitten der Reifezeugnisse. So geht man z.B. an der Universität München bei der Bewertung von 3 Fächergruppen aus, deren Durchschnittsnoten mit 3, 2 oder 1 multipliziert und addiert werden. Die Zulassungsgrenze lag im Sommersemester 1967 bei einer Bewertungszahl 19,41 (AUMILLER 1967). In Baden-Württemberg werden die Zensuren von 4 schriftlichen Prüfungsfächern im Abitur addiert, das beste Fach doppelt genommen und die Gesamtzahl durch 5 dividiert. In Kiel wurden Abiturzeugnisnoten im Zusammenhang mit Fragebogen zur Person oder persönlichem Vorstellen als Selektions-

modus gewählt. In Marburg erfolgte die Entscheidung über die Zulassung anhand der MP (s.o.). Die Kritik an einem sich an Schulleistungen orientierenden Zulassungssystem beginnt bei der Reflektion über den Aussagewert von Abiturzeugnisnoten. Daß die Abiturnoten keine objektive Aussage über den Wissensstand und die Leistungen eines Bewerbers zu geben vermögen, beweisen unter anderem KAPUSTE, NOACK, SCHÜTZ, WEINGARDT und BERNHART. Von KÜGELGEN und BARKMANN weisen darauf hin, daß die Benotung stark abhängig ist von der Individualität des Lehrers. FURCK äußert, daß der Lehrer bei der Beurteilung von Schülerleistungen keinen eindeutigen Maßstab besitzt. TESCHNER weist darauf hin, daß die gleichen Abiturnoten verschiedener Schularten unterschiedlich zu bewerten sind. KAPUSTE und SCHÜTZ bemerken die besonders regional abhängigen Schwankungen in der Güte der Abiturzeugnisse. Besonders energisch wendet sich von KÜGELGEN dagegen, daß "in Deutschland derzeit praktisch aufgrund einer zufällig bemessenen Nachbewertung der Abiturnoten (seitens der Universitäten) entschieden wird, wer Arzt werden darf" und ist sich mit KAPUSTE, GOLENHOFEN u.a. darin einig, daß der diagnostische und prognostische Aussagewert von Zeugnisnoten in bezug auf die Eignung zum Arztberuf sehr gering ist. Bedenklich stimmt die Auslese anhand von Schulnoten auch, wenn man die Ergebnisse von JUST (1939) bedenkt, der die Schulleistungen von Angehörigen bestimmter Berufe untersuchte und feststellte, daß Schulleistungen der Ärzte relativ niedrig liegen. Es ergibt sich daraus die Frage, ob bestimmte Berufe andere Aspekte der Intelligenz als die der "Schulintelligenz" erfordern. Über den Aussagewert der einzelnen Noten im Hinblick auf den späteren Studienerfolg zeigen die Untersuchungen von SCHÜTZ (1968), daß die in den MP enthaltenen Fächer Deutsch, Mathematik, Physik und Chemie nur unerhebliche Beziehungen zu den vorklinischen Prüfungen aufweisen. Zudem ist die Berechnung einer Kernfächerzensur für die Vorhersage ungünstiger als ein allgemeiner Abiturzeugnisnotendurchschnitt. Er kommt überdies zu dem Ergebnis, daß 70 bis 75 % der Bewerber mit guten Abiturnoten auch gute Prüfungsleistungen im Examen

bringen. Er fand andererseits, daß gute vorklinische Examina zu 50 bis 60 % von Studierenden mit Zeugnisdurchschnittsnoten von 3 und 4 erbracht wurden. Man sollte daher die Abiturnoten nicht überbewerten, allerdings auch nicht ganz außer acht lassen, sich jedoch bemühen, ihre Validität zu erhöhen (SCHÜTZ 1968). Erst dann könnten sie zu besseren Prädiktoren für den weiteren Studienerfolg werden (KAPUSTE 1966). Doch selbst dann wäre man dem erstrebenswerten Ziel, aus einer großen Zahl von Bewerbern nur die ungeeigneten zu selektieren, kaum näher gekommen. Die Garantie für bessere Studienleistungen würde sich zwar erhöhen, doch bleibt es eine offene Frage, ob sich Studenten mit guten vorklinischen Examina auch als gute Mediziner herausstellen. Denn daran und nur daran müßten sich die Aufnahmekriterien zum Studium der Medizin orientieren. Verschiedentlich bemüht man sich darum, dem gerechter zu werden. Von KÜGELGEN (1966) z.B. versucht in der Hauptsache von der Gesamtpersönlichkeit des Bewerbers her eine Entscheidung zu fällen. Den wissenschaftlichen Kriterien am ehesten stand hält jedoch nur das amerikanische Zulassungsverfahren, das bei der Bewerbung hauptsächlich die Ergebnisse von speziellen Eignungstests und Leistungstests berücksichtigt und den Schulnoten nur relativ geringe Bedeutung beimißt.

3) Klausuren

a) Multiple-Choice-Tests

Bei den vorliegenden Klausuren wurden, wie bereits oben erwähnt, zwei verschiedene Testarten verwandt. Einerseits verfuhr man nach der herkömmlichen Weise im Sinne von Fragen mit freiem Beantwortungsmodus, andererseits wurden Tests nach Multiple-Choice-Art durchgeführt. Diese auch als objektive Tests (HUBBARD 1968) bezeichneten Verfahren, die in Zukunft auch einen großen Teil der noch bisher mündlichen Examina ersetzen werden (GEORG 1970; SCHULZ, REMSCHMIDT, PRINZ 1970; KAPUSTE und NOACK 1968), sind auf der Grundlage der psychologischen Testtheorie und Testkonstruktion entwickelt worden (LIENERT 1967). Die

Ausweitung des Anwendungsbereichs (z.B. in der Medizin) erfolgte aus der Erkenntnis, daß herkömmliche schriftliche und mündliche Prüfungsverfahren wenig gültig und zuverlässig sind (DAHMER 1970). Besonders im amerikanischen Schrifttum liegen ausführliche Untersuchungen über die Vorteile dieses Testverfahrens vor. So findet sich unter anderem bei LEE und SYMONDS 1934; SYMONDS 1927; PATERSON 1926; EURICH 1961; BULL 1951; COWLESS und HUBBARD 1952; HUBBARD und CLEMANS 1960; u.a. übereinstimmend die Aussage, daß objektive Testverfahren weit aus reliabler sind und gewöhnlich auch bessere Validität haben als andere. Die Autoren PATERSON 1926; EURICH 1931; FASTIER 1959 und HUBBARD 1968 weisen außerdem darauf hin, daß diese Prüfungsart mehr als nur reines Faktenwissen prüfe und außerdem mehr mit der Intelligenz korreliere. Häufig werden darüber hinaus (SCHULZ, REMSCHMIDT, PRINZ 1970; KAPUSTE und NOACK 1968; DAHMER 1970; HÖHN 1969; u.a.) zeitökonomische Vorteile für diese Prüfungsart ins Feld geführt:

1. durch die Möglichkeit der gleichzeitigen Prüfung eines großen Kollektivs
2. durch die Möglichkeit, daß ein großes Wissensgebiet in kurzer Zeit detailliert geprüft werden kann
3. dadurch, daß die Auswahl der Fragen, sofern schon ein qualifizierter Item-Pool (SCHULZ und PRINZ 1969) vorhanden ist, sowie auch deren Auswertung per Computer erfolgen kann.

Diese genannten zeitlichen Vorteile erweisen sich aber dann als nur scheinbare, wenn man berücksichtigt, daß die laufende Kontrolle wie auch die ständige gewissenhafte Neuentwicklung und Erprobung der Items einen erheblichen Zeitaufwand bedeuten (HUBBARD 1968; INGENKAMP 1969). Zu den weiteren echten Vorteilen zählt die statistische Überprüfbarkeit der Items (GUILFORD 1956; LIENERT 1962). So kann deren Schwierigkeit eindeutig festgestellt werden, wie auch die Fähigkeit des Items zwischen guten und schlechten Testteilnehmern zu diskriminieren. Außerdem kann jede Alternative beurteilt werden, da ihre prozentuale Häufigkeitswahl und ihre Trennschärfe berechnet werden.

Die hier vorliegenden Tests erfüllen ihren Sinn insofern, als der Stoff des vorausgegangenen Semesters punktuell überprüft wird. Es ist jedoch einzuwenden, daß die Anzahl von 20 Fragen zu gering ist, um ein objektives Leistungsbild der Studenten in diesem Fach zu erhalten. Außerdem wurden die verwandten Fragen erstmalig und für dieses Tests geschaffen. Eine Vorprüfung der einzelnen Items mit ihren Alternativen mit nachfolgender statistischer Analyse auf Schwierigkeitsgrad und Trennschärfe ist nicht erfolgt. So konnte bis auf das Urteil der Dozenten auch keine Aussage gemacht werden, wie gut oder wie sinnvoll diese Tests sein würden. Weil zudem die prognostische Aussagefähigkeit dieser Tests für bestimmte Studienleistungen unbekannt war, scheint es zumindest problematisch, solche Tests darüber entscheiden zu lassen, wer Medizin studieren darf, abgesehen von der darüber hinausgehenden fraglichen Beziehung zwischen Studienleistung und Eignung für den Arztberuf.

b) Tests mit freier Beantwortung

Was oben für die Vorhersagefähigkeit für die Multiple-Choice-Tests gesagt wurde, gilt in dem selben Maße auch für die Tests mit freier Beantwortung. Es wird darüber hinaus diesem Testverfahren vorgeworfen, daß es ihm einer objektiven Auswertbarkeit mangle (HUBBARD 1968, u.a.). Dieser Vorwurf trifft auch auf die von uns untersuchten Prüfungen zu. So wurden für die Bewertung derselben Antworten von verschiedenen Prüfern bis zu einem Punkt Unterschied pro Aufgabe festgestellt. Vorteile dieser Testart sind dagegen, daß die Probanden nicht aus mehreren vorgegebenen Antwortmöglichkeiten sich der Richtigen erinnern, d.h. einen passiven Wissensschatz mobilisieren, sondern daß aktiv spontan verfügbares Wissen geprüft wird.

Die Chemieklausur nimmt unter allen Klausuren eine Sonderstellung ein durch die inhaltliche Homogenität des Testes. Die Physikklausur dagegen hat als entscheidenden Vorteil die große Stofffülle, die hauptsächlich "Basiswissen" prüft. Durch ihre Heterogenität wird eine objektivere Beurteilung des Gesamtwissensstandes in diesem Fach möglich. Der Proband hat

die Chance, Schwächen in einem Teilgebiet zu kompensieren. Im übrigen gelten auch für diese Klausuren die oben genannten allgemeinen Bedenken.

4) Diskussion des I-S-T-AMTHAUER

Der Intelligenz-Struktur-Test von AMTHAUER wurde herangezogen, um unabhängig von klausur-bezogenen Leistungskriterien zusätzlich eine Information über die Intelligenz und Eignung der Bewerber zu erhalten. Man war sich jedoch bewußt, daß die Ergebnisse dieses Intelligenztestes nur mit Einschränkungen zu verwerten seien. Der I-S-T stützt sich auf das theoretische Konzept der Mehrfaktorentheorie der Intelligenz nach THURSTONE (1940) und entspricht damit dem theoretischen Stand der 40er Jahre (MERZ). Forschungen, die sich mit der Intelligenz befassen und die verschiedenen Definitionen der Intelligenz sind fast Legion. CATTEL: "Man kann heute nur feststellen, daß wir von einer befriedigenden Theorie des mit Intelligenz intendierten Sachverhaltens noch weit entfernt sind" (JÄGER 1967). Somit sind auch Tests, die sich bemühen, die Intelligenz zu messen, nur von beschränktem Aussagewert für das einzelne Individuum, da sie nur einzelne Dimensionen der Intelligenz erfassen können (JÄGER 1967) und darüber hinaus das Ergebnis situationsabhängig ist, was besonders dann berücksichtigt werden sollte, wenn außer der allgemeinen Niveaubestimmung der Intelligenz besonderer Wert auf differenzielle Vorhersagen gelegt wird (GROFFMANN 1964). Wie sehr emotionale Faktoren bei der Intelligenzleistung eine Rolle spielen bringt ROHRACHER damit zum Ausdruck, daß er die Intelligenz definiert als den "durchschnittlichen Leistungsgrad der psychischen Funktionen bei der Bewältigung neuer Situationen unter normalen emotionalen Bedingungen" (LIENERT 1963). Empirisch verifizieren läßt sich außerdem nur ein kleiner Ausschnitt an Fähigkeitskonstrukten der Intelligenz (HEISS 1964). Die Beurteilung eines Probanden anhand eines standardisierten Meßwertes (IQ), der sich aus den Gesamtergebnissen berechnet, stellt sicherlich eine große Verallgemeinerung dar und läßt nur sehr begrenzte Aussagen über

die Funktion der so gemessenen Intelligenz zu (HEISS 1964). Die Profildarstellung der Intelligenz eines Bewerbers gibt schon eher Auskunft über bestimmte Schwerpunkte der Intelligenzleistungen. Ob jedoch eine Begabungsdiagnose aus dem Vergleich des Intelligenzprofils eines Studienbewerbers mit dem von AMTHAUER empirisch an nur 38 Medizinern ermittelten "Berufsprofil" gestellt werden kann scheint bedenklich, da weder eine wissenschaftlich begründete oder empirisch gewonnene Rechtfertigung dieses Vorgehens bekannt ist.

5) Diskussion der Vorphysikumsnoten

Bei der Verwendung dieser Noten als Ergebnis der ersten staatlichen Leistungskontrolle gilt es zu bedenken, daß auch sie kein objektives Maß des tatsächlichen Wissensstandes eines Studenten darstellen. Die entscheidenden Nachteile einer mündlichen Prüfung gegenüber einer objektiven Leistungskontrolle haben sicherlich nicht unerheblichen Einfluß auf die Notengebung. Die Nachteile mündlicher Prüfungen sind unter anderem (SCHULZ, REMSCHMIDT, PRINZ 1970):

1. psychologischer Art

- | | |
|-------------------------------------|---|
| a) vom Prüfer abhängige Variable: | α) individueller Bewertungsmaßstab |
| | β) unterschiedlicher Schwierigkeitsgrad der Fragen |
| | γ) Sympathie-Antipathie einem Prüfling gegenüber |
| b) vom Prüfling abhängige Variable: | α) der Wissensstand |
| | β) individuelle Formulierungsfähigkeit |
| c) Interpersonale Wechselwirkungen: | Prüfungsummosphäre |

2. Als prüfungstechnische Nachteile haben sich herausgestellt:

- a) unterschiedliche Anforderungen an die Studenten
- b) fachlich punktuelle Prüfungen
- c) unterschiedliche Bewertungsmaßstäbe
- d) Nichtüberprüfbarkeit der Bewertungen.

Darüber hinaus besteht die Gefahr, daß nicht die einzelne Leistung, sondern der Gesamtablauf der Prüfungsgespräche beurteilt wird.

Der offensichtlichen Nachteil dieser mündlichen Prüfungen bewußt, wurden die Ergebnisse des Vorphysikums trotzdem als erster Indikator für den Studienerfolg verwandt, nicht zuletzt um auch Vergleiche mit anderen wissenschaftlichen Untersuchungen, die dieselben Prüfungen berücksichtigen, anstellen zu können. Im übrigen gelten für die Zensuren dieser Prüfungen das gleiche, was oben bereits über den prognostischen Wert von Abiturnoten bezüglich der spezifischen Eignung zum Arztberuf gesagt wurde.

V. Statistische Auswertungsmethoden

1. Als grundlegende statistische Kennwerte wurden Mittelwerte (MW), Varianzen und Standardabweichungen berechnet.

2. Prüfung der Normalverteilung:

Die Prüfung auf Normalverteilung der Variablen wurde mit Hilfe des bekannten Tests, der Schiefe und Exzess berücksichtigt, durchgeführt (SACHS).

3. Die statistische Berechnung von Zusammenhängen zwischen Variablen erfolgt mittels des Korrelationskoeffizienten (r). Sein Wert berechnet sich nach der Formel:

$$r_{xy} = S_{xy} / \sqrt{S_{xx} \cdot S_{yy}} = s_{xy} / s_x \cdot s_y \quad (\text{Formel 1})$$

Es bedeutet:

- S_{xx} = Summe der Abweichungsquadrate der Variablen x
- S_{yy} = Summe der Abweichungsquadrate der Variablen y
- S_{xy} = Summe der Abweichungsprodukte der Variablen x und y
- s_{xy} = Kovarianz
- s_x = Standardabweichung von x
- s_y = Standardabweichung von y

Der Korrelationskoeffizient r gibt Auskunft über die Enge des linearen Zusammenhangs jeweils zweier Variablen. Seine Werte liegen zwischen -1 und +1. Keine lineare Beziehung ist bei $r = 0$ vorhanden. Das Vorzeichen entspricht der Steigung der Regressionsgeraden (ÜBERLA 1968). Die obige Formel ist der sogenannte "Produkt-Moment-Korrelationskoeffizient". Er ist bei quantitativen Merkmalen mit linearer Abhängigkeit zu verwenden (MITTENECKER 1962).

Die Wechselbeziehungen zwischen Merkmalen, von denen das eine quantitativ (z.B. Testpunktwerte), das andere echt alternativ (z.B. richtige oder falsche Lösung) gemessen werden kann, werden mit dem punktbiserialen Korrelationskoeffizienten berechnet.

Seine Formel ist:

$$r_{pbis} = \frac{\bar{x}_1 - \bar{x}_0}{s_x} \cdot \sqrt{p(1-p)} \quad (\text{Formel 2})$$

(SCHULZ, REMSCHMIDT, PRINZ 1970; LIENERT 1968).

Es ist:

- $\bar{x}_1$ = MW der Rohwerte der Pbn mit richtiger Lösung
- $\bar{x}_0$ = MW der Pbn mit falscher Lösung
- s_x = Standardabweichung der Rohwerte
- $p = \frac{N_r}{N}$ Bruchteil der Individuen mit richtiger Lösung

r_{pbis} ist ein Produkt-Moment-Korrelationskoeffizient, wenn man die Alternativmerkmale "richtig" oder "falsch" mit 0 oder 1 bewertet. Algebraisch sind beide Koeffizienten identisch (LIENERT 1967). Sind die beiden zu korrelierenden Merkmale echt alternativ (beide Variable nehmen nur die Werte 0 und 1 an), so wird der Zusammenhang dieser beiden Variablen durch den Phi-Koeffizienten gemessen:

$$\Phi = \frac{p_{ij} - p_i p_j}{\sqrt{p_i (1-p_i)} \cdot \sqrt{p_j (1-p_j)}} \quad (\text{Formel 3})$$

p_{ij} ist die Häufigkeit, mit der Aufgabe i und j von den Probanden gleichzeitig richtig gelöst wurden,

p_i ist die Häufigkeit, mit der die Aufgabe i richtig gelöst wurde.

Φ ist ebenfalls ein Produkt-Moment-Korrelationskoeffizient.

Aus der herkömmlichen Statistik ist bekannt, daß der Korrelationskoeffizient bei gewissen Voraussetzungen über Normalverteilungseigenschaften der Variablen x und y mit Hilfe der FISCHER-Formel auf Abweichung von 0 getestet werden kann.

$$t = \frac{r}{\sqrt{1 - r^2}} \cdot \sqrt{N - 2} \quad (\text{Formel 4})$$

Untersuchungen von KENDALL und STUART zeigen aber, daß dieser t-Test robust gegenüber Abweichung von der Normalverteilungsvoraussetzung ist. Deshalb gilt er auch approximativ für Korrelationen zwischen beliebigen quantitativen Variablen. Außerdem ist er anwendbar für Korrelationen zwischen einer quantitativen und einer alternativen Variablen (LIENERT, KENDALL und STUART). Sind beide Variable alternativ, so wird die signifikante Abweichung von Null mit Hilfe der χ^2 -Verteilung getestet. LIENERT führt an, daß die Beziehung gilt:

$$\mathfrak{I} = \sqrt{\frac{X_1^2}{N}} \quad (\text{Formel 5}) \quad (N = \text{Stichprobenumfang})$$

Hierbei berechnet sich X^2 nach dem üblichen Vier-Felder-Schema. Zur Überprüfung der MW-Unterschiede in den Leistungen des I-S-T der nach über- und unterdurchschnittlichen Ergebnissen in den Klausuren bzw. Vorphysikum in 2 Gruppen geteilten Testteilnehmer wurde der parameterfreie Test von MANN und WITNEY (SACHS) durchgeführt.

Im Rahmen der Testanalyse kann man mit Hilfe des Korrelationskoeffizienten wesentliche Aussagen machen über die "Gütekriterien" eines Tests und seiner Items (SCHULZ, REMSCHMIDT, PRINZ 1970). Es sind dies:

a) Die Validität

Die Validität ist ein Maß für die "diagnostische Verlässlichkeit" eines Testes, d.h. der Validitätskoeffizient gibt an, inwieweit der Test tatsächlich das mißt, was er messen soll. Man ermittelt diesen Koeffizienten, indem das Testergebnis mit einem Außenkriterium, das die gleichen Eigenschaften zu messen beansprucht, korreliert wird. Da es sich hier um einen reinen Produkt-Moment-Korrelationskoeffizienten handelt, kann er mit der FISCHER-Formel (s.o) auf Signifikanz getestet werden. Das Außenkriterium, mit dem der Test korreliert wird, muß vorher festgelegt sein. Es kann dieses z.B. eine auf andere Weise ermittelte "Eignung für den medizinischen Beruf" sein, oder "Wissensstand im Fach...". Man erhält dann eine sogenannte "kriterienbezogene Validität", die definiert wird durch den Bezug der Testpunktwerte zu den Kriterienpunktwerten (LIENERT 1967). Die Validität unserer Leistungstests kann in Ermangelung von guten Validitätskriterien nicht optimal ermittelt werden, denn einerseits ist ein nicht anfechtbares Verfahren zur Objektivierung der Eignung für den medizinischen Beruf nicht bekannt, andererseits stand uns als einziges zu korrelierendes Maß für den Wissensstand in den jeweiligen Fächern nur das Ergebnis der mündlichen Vorphysikumsprüfung

zur Verfügung, das aus obengenannten Gründen kein objektives Leistungsmaß darstellt.

b) Die Reliabilität

Der Reliabilitätskoeffizient ist Ausdruck für die statistische Zuverlässigkeit eines Testes, für "den Grad der Genauigkeit, mit der er ein bestimmtes Persönlichkeits- oder Verhaltensmerkmal mißt" (LIENERT 1968). Die Genauigkeit bezieht sich dabei nur auf den statistischen Meßwert und kennzeichnet die Leistung des Testes als Meßinstrument. Es gibt verschiedene Methoden, diesen Koeffizienten zu ermitteln. In unserem Fall wurde er berechnet nach der Formel 20 von KUDER-RICHARDSON:

$$r_{tt} = \frac{n}{n-1} \cdot \left(1 - \frac{\sum p_i(1-p_i)}{s_x^2} \right) \quad (\text{Formel 6})$$

p_i = Häufigkeit der Wahl der richtigen Alternative.

Bei Tests, die ein inkonstantes Merkmal (z.B. Wissen in...) messen, wird nach obiger Formel die statistische Zuverlässigkeit als "innere Konsistenz" gemessen. Es zeigt sich dabei, daß heterogene Tests gewöhnlich sehr viel niedrigere innere Konsistenz aufweisen, als homogene. Dafür korrelieren sie widerum höher mit den "praktischen Validitätskriterien" (LIENERT 1968).

Bei den Tests mit freier Beantwortung wird der Reliabilitätskoeffizient nach der Split-Half-Methode geschätzt, indem man die Aufgaben in zwei Gruppen aufteilt, für jeden Probanden die Rohwerte x_1 und x_2 in den beiden Testhälften bestimmt und die Korrelation $r_{x_1x_2}$ berechnet. Da diese Korrelation nur die Zuverlässigkeit für den Test mit der Hälfte der Aufgaben angibt, berechnet man r_{tt} nach SPEARMAN-BROWN durch

$$r_{tt} = \frac{2 r_{x_1x_2}}{1 + r_{x_1x_2}} \quad (\text{Formel 7}).$$

c) Der Trennschärfeindex (TI)

Die detailliertere Analyse der einzelnen Testitems erfolgt mittels eines speziellen punktbiserialen Korrelationskoeffizienten zwischen der Leistung in einer Aufgabe und dem Gesamtwert des Testes. Der Trennschärfeindex kennzeichnet die diskriminierende Funktion der Aufgabe, da sein Wert ein Maß ist für die Eigenschaft des Items, zwischen Pbn mit guter und schlechter Testgesamtleistung zu differenzieren (SCHRADER 1967).

Wie trennscharf jede Alternative bei Multiple-Choice-Aufgaben ist, wird nach der Formel 2 als "Leistungskorrelation" berechnet. Die Wahl eines Ablenkers wird mit 1, seine Ablehnung mit 0 bewertet und mit dem Rohwert nach der Formel für den r_{pbis} korreliert. Der Wert der Leistungskorrelation soll bei gut konstruierten Multiple-Choice-Aufgaben nur bei der richtigen Alternative positiv sein, bei den "distractors" dagegen negativ oder annähernd Null sein, was bedeutet, daß hauptsächlich "schlechte" Pbn sich für die falschen Lösungen entscheiden.

d) Der Schwierigkeitsindex (SI)

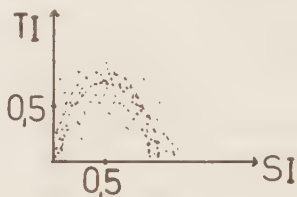
Die einzelnen Testitems werden auf ihre Schwierigkeit überprüft, indem man nach

$$SI = \frac{N_r}{N} \cdot 100$$

$$\begin{aligned} N_r &= \text{Summe der Pbn mit richtiger Antwort} \\ N &= \text{Summe aller Antworten} \end{aligned}$$

einen Schwierigkeitsindex berechnet. Er ist definiert als der "prozentuale Anteil p der auf dieser Aufgabe entfallenden richtigen Antworten in Beziehung zur Analysenstichprobe N" (LIENERT 1968).

Die Beziehung der beiden Größen SI und TI wird nach LIENERT durch folgende Parabel charakterisiert:



Optimal trennscharf sind danach Aufgaben mittlerer Schwierigkeit. Sehr schwere und sehr leichte Aufgaben trennen das Kollektiv dagegen nur schlecht.

Entsprechende Kenngrößen bei den Free-Response-Tests:

Die Schwierigkeit der Aufgabe läßt sich hier berechnen, indem man den Quotienten aus mittlerer Punktzahl $\bar{x}_i$ und maximal erreichbarer Punktzahl ($x_{\max_i}$) mit 100 multipliziert:

$$SI = \frac{\bar{x}_i}{x_{\max_i}}$$

($\bar{x}_i$ = MW der erreichten Punkte beim Subtest i).

Die Trennschärfe wird mit Hilfe der Produkt-Moment-Korrelation zwischen den erreichten Punkten bei der Aufgabe i (x_i) und dem Gesamtrahwert (x) gemessen: $r = r_{x, x_i}$.

4. Die Faktorenanalyse

Das statistische Verfahren der Faktorenanalyse (FA) wurde hauptsächlich durch SPEARMAN (1863 - 1945) und THURSTONE (1887 - 1955) entwickelt (HOFSTÄTTER 1966). Sie stellt eine Analyse der Wechselbeziehungen zwischen Variablen dar. Sie geht daher von der Korrelationsmatrix aus, die die Information über das "Zueinander" aller zu untersuchenden Variablen eines Kollektivs enthält (ÜBERLA 1968). Aus den Interkorrelationen der einzelnen Variablen lassen sich Faktoren berechnen (auch Dimensionen, Grundfunktionen, Grundvariable bezeichnet. PAWLIK 1968), die am Zustandekommen der Leistungen und ihrer sich gegenseitig

bedingenden Wechselbeziehungen beteiligt sind. Es wird dargelegt, welche Faktoren auf die einzelnen Variablen einwirken und inwieweit die Variablen mit bestimmten Faktoren geladen sind (HEHLMANN 1965). Bei diesen Faktoren handelt es sich um "hypothetische, nicht meßbare Einflußgrößen", die jedoch die Zusammenhänge erklären können (Erklärung in formalmathematischem Sinn), außerdem eine Reduktion der Datenmenge sowie eine möglichst genaue Reproduktion der Variablenbeziehungen ermöglichen (ÜBERLA 1968). Das statistische Verfahren der FA hat somit ordnende wie auch hypothesenbildende Funktion. Die Faktoren sind dabei graduell unterschiedlich an der Varianz σ^2 einer Variablen beteiligt (JÄGER 1967).

Die Summe der Quadrate aller Faktorenladungen ist 1. die Einheitsvarianz σ^2 einer Variablen setzt sich zusammen aus:
 $h^2 + u^2 = \sigma^2 = 1$, wobei h^2 die Summe der Quadrate der gemeinsamen Faktoren (common factors) ist und als Kommunalität bezeichnet wird $h^2 = a_{11}^2 + \dots a_{ir}^2$.

(a_{11}^2 = Varianzanteil des 1. Common-Factors)

Über die Einzelvarianzen u^2 ($u^2 = 1 - h^2$) kann die FA nichts aussagen. u^2 enthält jene durch die Spezifität der Variablen bedingten Varianzanteile, sowie die durch Irrtum verursachten Fehlervarianzen. Durch die faktorenanalytische Lösung werden die Kommunalitäten der Variablen erklärt (JÄGER 1967), d.h. die Varianzanteile eines Merkmals werden in gemeinsame Faktoren und Restvarianz zerlegt (PAWLIK 1968). Die Varianzanteile jedoch, deren Bedingungen durch die FA aufgeklärt werden sollen, kommen in den Interkorrelationen der Variablen zum Ausdruck. Die Korrelationen werden auf Faktoren reduziert und auf diese gemeinsamen Grundfunktionen zurückgeführt. Man kann die Korrelation r_{xy} zweier Variablen als Skalarprodukt von n gemeinsamen Faktorenladungen darstellen:

$$r_{xy} = a_{1x} \cdot a_{1y} + a_{2x} \cdot a_{2y} + \dots a_{nx} \cdot a_{ny}$$

a_{1x} = Ladung des 1. Faktors in Variable x

a_{1y} = Ladung des 1. Faktors in Variable y (JÄGER 1967).

Die FA ist im wesentlichen ein deskriptives Verfahren. Das mathematisch-statistische Verfahren liefert Faktoren, die der Interpretation zugänglich sind. Die Interpretation erfolgt nur aus den Gemeinsamkeiten der Variablen, die mit einem Faktor besonders hoch geladen sind (HEHLMANN 1965).

5. Bei uns wurden hauptsächlich Hauptkomponentenanalysen (HKA) durchgeführt. Beide Verfahren gegeneinander abzugrenzen ist teilweise eine Nomenklaturfrage (ÜBERLA 1968). Die HKA wird besonders dann als beschreibendes Verfahren angesetzt, wenn zunächst viele Daten auf einige wenige Variable reduziert werden sollen und dabei die zu wählende Anzahl der reduzierten Variablen (Dimensionen) unbekannt ist. Geometrisch veranschaulicht bilden die Meßwerte der einzelnen Pbn häufig ellipsoidförmige Punktwolken in einem mehrdimensionalen Raum (ÜBERLA 1968). Mit der Hauptachsenmethode wird zunächst die längste Achse dieses Ellipsoids bestimmt, dann die zweitlängste, die senkrecht auf der ersten steht, die drittlängste, die orthogonal zur ersten und zweiten ist usw. bis alle Hauptachsen nacheinander bestimmt sind. Die Länge der jeweiligen Hauptachse entspricht der Größe der Varianz in ihrer Richtung. Da die Gesamtvarianz bekannt ist, läßt sich ermitteln, wieviel Prozent der Gesamtvarianz in einer Hyperebene von r Hauptachsen enthalten ist. Die Berechnung der Hauptachsen geht von der Korrelationsmatrix aus. Jeder Hauptachse entspricht ein Eigenvektor α und ein Eigenwert λ . (ÜBERLA 1968). Der Eigenwert ist gleich der Varianz in Richtung der Hauptachse, dessen Länge der Wurzel von λ entspricht.

Nach Inspektion der Eigenwerte und Eigenvektoren wird die Zahl der Dimensionen festgelegt. Als Extraktionskriterium gilt: a) nach dem "Scree-Test" (CATTELL 1966) Knick im Verlauf der Eigenwerte, b) 70 % aufgeklärte Varianz, bes. für allgemeine Untersuchungen. Nach Selektion der Dimensionen wird dann zur besseren Interpretation eine Rotation auf Einfachstruktur nach dem Varimaxkriterium durchgeführt. Berücksichtigung bei der Interpretation finden hauptsächlich die größten Eigenwerte, da sie die meiste Varianz aufklären. Bei heterogenen Variablen kann es dabei vorkommen, daß nur wenig Varianz sinnvoll interpretiert werden kann.

VI. Darstellung der Ergebnisse

Alle Berechnungen wurden mit Bibliotheksprogrammen des Instituts für medizinisch-biologische Statistik und Dokumentation der Universität Marburg durchgeführt.

Abb. 1

1. Klausurstatistiken

	BOTANIK	ZOOLOGIE	PHYSIK	CHEMIE
a) Teilnehmerzahl	84	84	84	84
b) Anzahl der Aufgaben	20	20	10	10
c) Erreichbare Punkte	20	20	100	200
d) Mittelwert	8,30	9,98	49,0	69,2
e) Streuung	2,68	3,10	21,0	18,8
f) Reliabilität	0,45	0,57	0,999**	0,998**
g)* Validität	0,24	0,12	0,33	0,03
h) Variationsbreite der SI	4,76 - 82,1	7,14 - 78,5	29 - 72,3	15 - 49,7
i) MW der SI	36,2	48,87	56,23	34,9
j) Variationsbreite der TI	0,08 - 0,47	0,07 - 0,53	0,06 - 0,76	0,49 - 0,76

* Die Validierung der Klausuren erfolgte durch Korrelierung mit den entsprechenden fachlichen Leistungen des Vorphysikums. Dieser Koeffizient wurde daher an 36 Pbn ermittelt. Alle 4 Validitätskoeffizienten sind nicht signifikant, da nach Formel 4 nur Koeffizienten größer als 0,34 signifikant auf dem 5 % - Niveau sind.

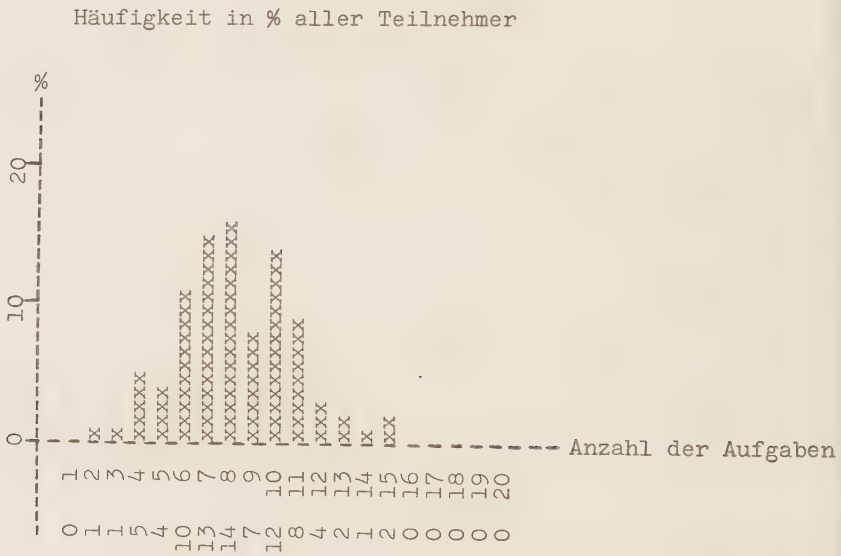
** Diese Koeffizienten wurden nach der Split-half-Methode berechnet.

Abb. 2

2. Rohwerteverteilungen

Rohwerteverteilungen der Multiple-Choice-Klausuren

a). BOTANIK



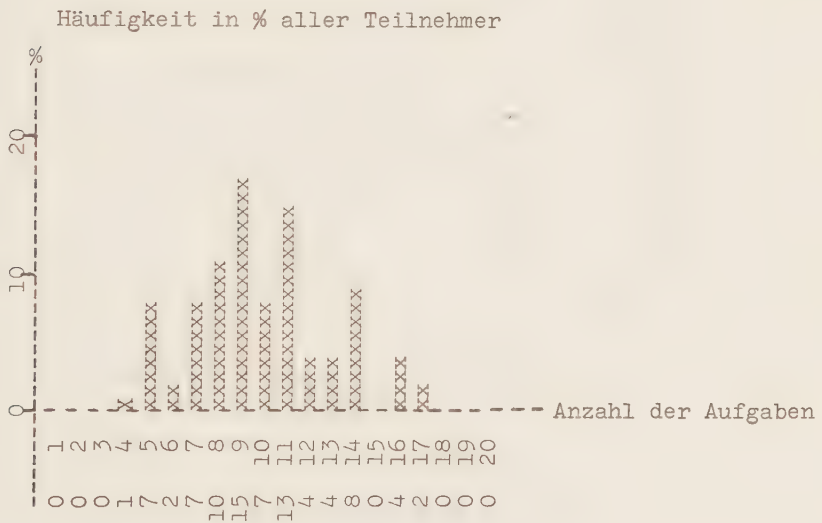
Die Überprüfung der Werte auf Normalverteilung ergab, daß die Ergebnisse der Botanikklausur normalverteilt sind.

Abb. 3

2. Rohwerteverteilungen

Rohwerteverteilungen der Multiple-Choice-Klausuren

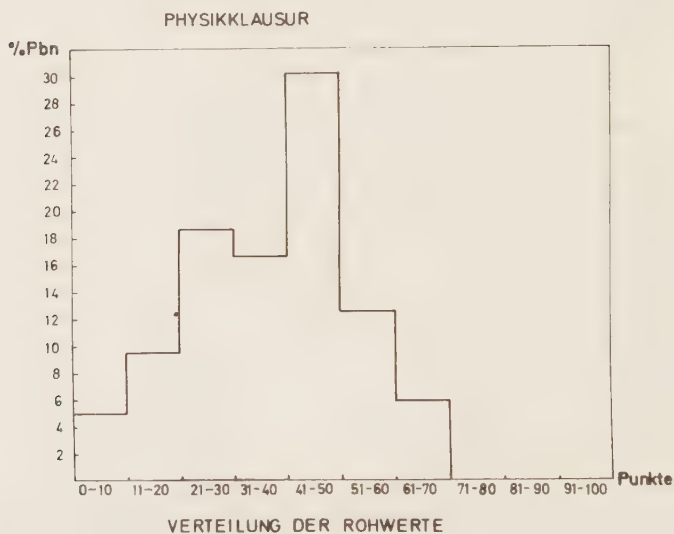
b) ZOOLOGIE



Die Berechnung ergab Normalverteilung.

Abb. 4

c) Rohwerteverteilung der Physiklausur. Klassengrenzen in erreichten Punktzahlen angeben.



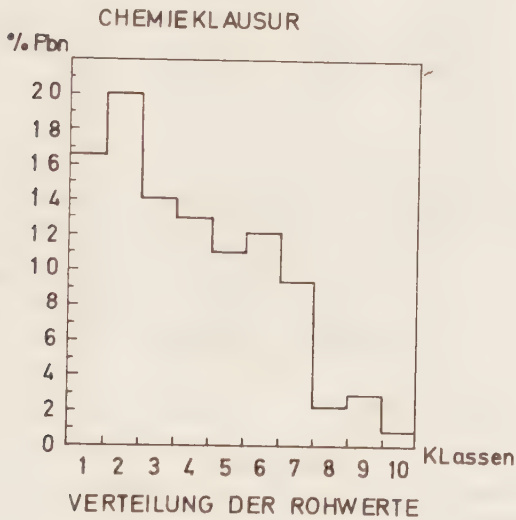
Die Berechnung ergab Normalverteilung.

Abb. 5

d) Rohwerteverteilung der Chemieklausur.

Klasseneinteilung der Punktwerte:

Kl. 1 = 0,00 - 19,9;	Kl. 2 = 20,0 - 39,9
Kl. 3 = 40,0 - 59,9;	Kl. 4 = 60,0 - 79,9
Kl. 5 = 80,0 - 99,9;	Kl. 6 = 100,0 - 119,9
Kl. 7 = 120,0 - 139,9;	Kl. 8 = 140,0 - 159,9
Kl. 9 = 160,0 - 179,9;	Kl. 10 = 180,0 - 200,0



Die Berechnung ergab auf dem 1 % - Signifikanzniveau Normalverteilung. Auf höherem Niveau jedoch nicht mehr.

3. Vergleich der SI (Häufigkeitsverteilungen)

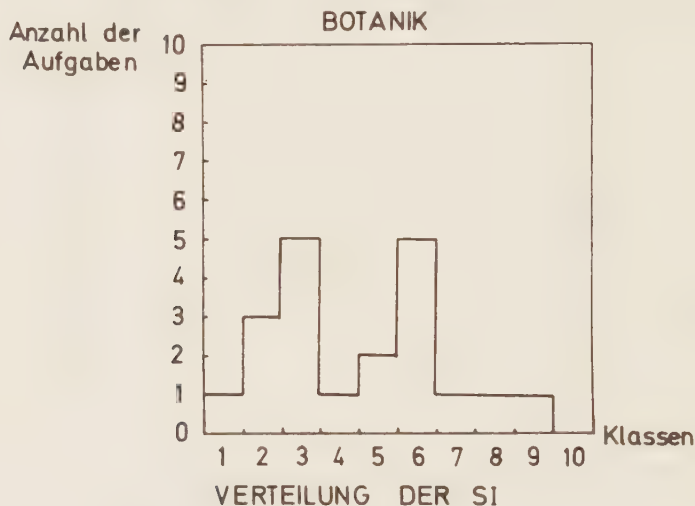
1. Multiple-Choice-Aufgaben

(20 Items pro Test)

Klasseneinteilung siehe Abb. 8

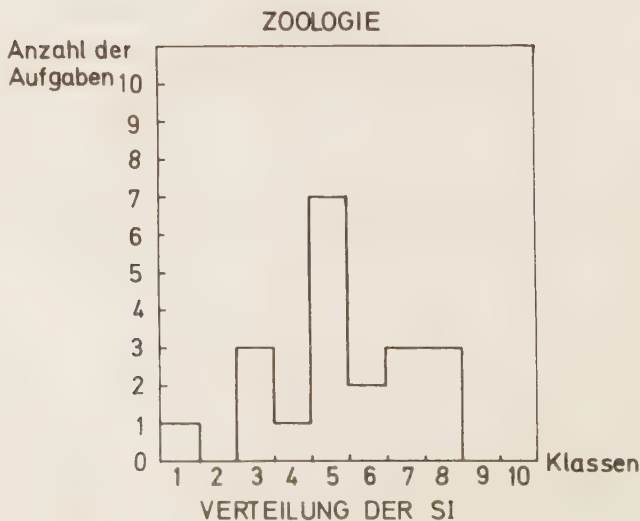
Abb. 6

a) BOTANIK



b) ZOOLOGIE

Abb. 7



Aufgaben der Klausuren mit freier Beantwortung
(10 Aufgaben pro Test)

Abb. 8

Physik und Chemie

Klassen Nr.	Schwierigkeit in %	<u>Häufigkeiten</u>	
		PHY	CHE
1.	0,00 - 9,9	0	0
2.	10,0 - 19,9	0	1
3.	20,0 - 29,9	1	2
4.	30,0 - 39,9	1	3
5.	40,0 - 49,9	2	4
6.	50,0 - 59,9	1	0
7.	60,0 - 69,9	3	0
8.	70,0 - 79,9	2	0
9.	80,0 - 90,0	0	0

Abb. 9

4. Vergleich der Trennschärfeindices

Nr.	Schwierig- keit in % Klassen- grenzen	Klausuren mit 20 Items		Klausuren mit 10 Items	
		BO	ZOO	PHY	CHE
1.	0,00-0,09	1	2	1	0
2.	0,10-0,19	4	1	0	0
3.	0,20-0,29	5	6	1	0
4.	0,30-0,39	7	7	0	0
5.	0,40-0,49	3	2	1	2
6.	0,50-0,59	0	2	6	2
7.	0,60-0,69	0	0	0	3
8.	0,70-0,79	0	0	1	3
9.	0,80-0,89	0	0	0	0

Abb. 11

b) Zoologie

Nr.	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0,25			0,27													
2			0,25		0,27	0,30				0,26							
3			0,24											0,29			
4					0,25					0,26							
5																	
6						0,26				0,25						0,25	
7										0,40							
8						0,30							0,34	0,26	0,26		0,28
9														-0,25			
10								-0,30									
12																	
14											0,24		-0,25	0,36			

Iteminterkorrelationsmuster der Zoologieklausur. Koeffizienten und Signifikanz wie in Abb. 10.

Matrix der Iteminterkorrelationen der Klausuren mit freier
Beantwortung

Abb. 12

Chemieklausur

1	2	3	4	5	6	7	8	9	10	Nr.
	.52	.62	.55	.48	.44	.88	.39	.42	.52	1
		.59	.57	.63	.40	.53	.36	.45	.56	2
			.47	.50	.49	.43	.28	.39	.52	3
				.42	.40	.28	.27	.52	.48	4
					.32	.50	.39	.32	.41	5
						.22	.48	.44	.47	6
							.22	.34	.35	7
								.22	.45	8
									.39	9
										10

Die Signifikanz wurde über t (FISCHER-Formel) geprüft. Bei $\alpha = 5 \%$ ist $r = 0,214$ signifikant.

Abb. 13

Physikklausur - Iteminterkorrelationsmuster

1	2	3	4	5	6	7	8	9	10	Nr.
				.29	.23	.38	.37	.22	.40	1
										2
						.34	.22	.23		3
						.26		.40	.30	4
				.38	.22	.18	.33	.39		5
					.23	.25	.38	.38		6
							.30	.28		7
										8
										9
										10

Die Signifikanz wurde über t (FISCHER-Formel) geprüft. Bei $\alpha = 5\%$ ist $r = 0,214$ signifikant.

6. Interne HKA der Klausuren

Interne Faktorenanalyse der Multiple-choice-Klausuren

a) Botanik

Abb. 14

Eigenwerte der Faktoren:

F 1	F 2	F 3	F 4	F 5	F 6	F 7	F 8	F 9
2,42	1,97	1,82	1,60	1,57	1,45	1,16	1,08	0,94

Prozentuale Beteiligung der 3 Faktoren an der
Aufklärung der Varianz:

F 1	F 2	F 3	F 4	F 5	F 6	F 7	F 8	F 9
12 %	10 %	10 %	7 %	7 %	8 %	6 %	5 %	5 %

Es wurden 3 Faktoren extrahiert, rotiert und interpretiert. Als Extraktionskriterium wurde nach dem "Scree-Test" von CATTELL (1966) der Knick im Verlauf der Eigenwerte verwandt.

Faktorenloadungen:

Item-Nr.	F 1	F 2	F 3
1	-0,01	0,62	0,14
2	0,08	0,22	0,00
3	0,29	-0,03	-0,13
4	-0,05	-0,01	0,49
5	0,40	-0,20	-0,01
6	-0,03	0,54	-0,10
7	0,58	0,10	-0,02
8	0,50	0,06	-0,04
9	-0,2	-0,27	0,72
10	0,24	0,21	0,52
11	0,17	0,47	0,63
12	0,34	-0,44	-0,30
13	0,04	0,46	-0,05
14	-0,00	-0,34	-0,46
15	0,28	-0,50	0,26
16	0,36	-0,18	0,06
17	0,32	0,23	-0,18
18	0,37	0,14	0,14
19	0,57	0,07	0,16
20	0,50	-0,07	0,45

Faktorenmuster:

Nr.	F 1	F 2	F 3
1		x	
2			
3			
4			x
5	x		
6		x	
7	x		
8	x		
9			x
10			x
11		x	x
12		-x	
13		x	
14			x
15		-x	
16			
17			
18			
19	x		
20	x		x

Berücksichtigt wurden nur Ladungen $\geq 0,4$. F = Faktor.

Interne Faktorenanalyse der Multiple-Choice-Klausur in Zoologie:

Abb. 15

Eigenwerte der ersten 8 Faktoren:

F 1	F 2	F 3	F 4	F 5	F 6	F 7	F 8
2,84	2,14	1,82	1,58	1,53	1,31	1,12	1,02

Prozentualer Anteil der durch die 3 Faktoren aufgeklärten Varianz:

F 1	F 2	F 3	F 4	F 5	F 6	F 7	F 8
15 %	10 %	9 %	6 %	8 %	6 %	6 %	5 %

Es wurden 3 Faktoren extrahiert, rotiert und interpretiert.
Extraktionskriterium nach dem "Scree-Test" von CATTELL (1966),
Knick im Verlauf der Eigenwerte.

Faktorenladungen:

Faktorenmuster:

Item-Nr.	F 1	F 2	F 3	Nr.	F 1	F 2	F 3
1	0,02	-0,49	0,45	1		-x	x
2	0,62	0,14	-0,10	2	x		
3	0,48	0,49	-0,18	3	x	x	
4	0,33	0,02	0,50	4			x
5	0,07	-0,04	0,03	5			
6	0,50	0,19	0,05	6	x		
7	0,11	-0,61	-0,03	7		-x	
8	0,46	-0,25	0,23	8	x		
9	0,68	-0,30	-0,05	9	x		
10	0,24	0,13	0,40	10			x
11	-0,11	0,12	-0,38	11			
12	0,46	0,23	0,26	12	x		
13	0,54	-0,26	0,05	13	x		
14	-0,03	0,23	0,67	14			x
15	0,02	0,08	-0,52	15			-x
16	0,40	-0,31	-0,55	16			-x
17	0,09	-0,71	-0,10	17		-x	
18	0,35	-0,06	0,03	18			
19	0,02	0,38	0,06	19			
20	0,44	0,03	0,25	20	x		

Berücksichtigt wurden hier Ladungen $\geq 0,4$.

Faktorenmuster der internen Faktorenanalysen

Die Faktorenanalysen wurden nach der Hauptkomponentenmethode mit anschließender Varimax-Rotation durchgeführt.

Es wurden jeweils 3 Faktoren extrahiert, um die inhaltliche Gliederung der Tests zu überprüfen.

a) Physikklausur

Abb.16

Eigenwerte der ersten 5 Faktoren:

3,02 1,27 1,11 1,04 0,96

Prozentualer Anteil der aufgeklärten Varianz:

30 % 13 % 11 % 10 % 10 %

Extraktionskriterium: Knick der Eigenwerte nach CATTEL (1966).
3 Faktoren extrahiert, rotiert und interpretiert.

Faktorenladungen:

Faktorenmuster:

Item-Nr. Faktor							
	I	II	III	Nr.	F 1	F 2	F 3
1	0,76	0,11	0,12	1	x		
2	0,21	0,17	0,73	2			x
3	0,25	0,83	0,19	3		x	
4	0,68	0,13	0,20	4	x		
5	0,46	0,21	0,47	5			
6	0,42	0,41	0,30	6			x
7	0,34	0,67	0,13	7		x	
8	0,25	0,23	0,55	8			x
9	0,38	0,55	0,02	9		x	
10	0,72	0,07	0,13	10	x		

Berücksichtigt wurden Ladungen $\geq 0,5$.

Interne FA der Chemieklausur

Abb. 17

	F 1	F 2	F 3
Eigenwerte	4,96	1,03	0,76
Prozentualer Anteil der durch die 3 Faktoren aufgeklärten Varianz	50 %	10 %	8 %

Extraktionskriterium anhand der Eigenwerte nach CATTELL (1966).

3 Faktoren extrahiert, rotiert, interpretiert.

Var. Nr.	F 1	F 2	F 3
1	0,48	0,31	0,52
2	0,69	0,22	0,43
3	0,57	0,23	0,49
4	0,27	0,22	0,73
5	0,78	0,30	0,12
6	0,09	0,66	0,47
7	0,82	0,00	0,14
8	0,18	0,88	0,08
9	0,14	0,13	0,80
10	0,38	0,52	0,40

Nr.	F 1	F 2	F 3
1			x
2	x		
3	x		
4			x
5	x		
6		x	
7	x		
8		x	
9			x
10		x	

Berücksichtigt wurden
Ladungen $\geq 0,5$

Abb. 18

7. Korrelationen zwischen Klausurergebnissen und Abiturnoten

84 Pbn

	Bot.	Chemie	Zool.	Physik	MP	Deutsch	Mathem.	Gemk.	Physik	Chemie
Botanik	1	0,54	0,33	0,40	0,34	-0,26	-0,39	-	-	-
Chemie		1	0,58	0,52	0,43	-	-0,45	-0,28	-	-
Zoologie			1	0,47	0,31	-	-0,35	-	-	-
Physik				1	0,32	-	-0,26	-	-	-
MP					1	-0,35	-0,62	-	-0,59	-0,61
Deutsch						1	-	0,26	-	-
Mathem.							1	-	0,26	0,36
Gem.Kunde								1	-	-
Physik									1	0,30
Chemie										1

Nach t sind Korrelationen von $r \geq 0,26$ signifikant.

8) Interkorrelationen der Intelligenztestergebnisse und Korrelationen der I-S-T-AMTHAUER-Items mit den Leistungen in den 4 Klausuren:

Abb. 19

	SE	WA	AN	GE	ME	RA	ZR	FA	WÜ	IQ	BO	CHE	ZOOL	PH
SE	1	0,37	-	0,32	-	0,26	-	-	-	0,43	-	-	-	-
WA		1	0,32	0,48	-	-	0,29	0,53	-	0,68	-	-	-	-
AN			1	-	-	0,41	0,30	-	-	0,51	-	-	-	-
GE				1	-	0,37	0,37	0,38	-	0,68	0,38	-	-	0,2
ME					1	-	-	-	-	0,39	-	0,33	0,27	-
RA						1	0,34	0,32	-	0,60	-	-	-	0,3
ZR							1	-	-	0,54	0,36	-	0,26	-
FA								1	-	0,66	0,31	-	-	-
WÜ									1	0,27	-	-	-	-
IQ										1	0,42	0,30	0,26	0

$r \geq 0,26$ ist signifikant nach t

Folgende Korrelationen bestehen zwischen dem I-S-T-AMTHAUER und dem Abitur:

Es korreliert Mathematik mit "GE" zu $r = 0,28$
mit "ZR" zu $r = 0,38$

Gemeinschaftskunde mit "ME" zu $r = 0,28$
mit "WÜ" zu $r = 0,26$.

Darüber hinaus liegen keine signifikanten Korrelationen vor.

10) Mittelwerte und Streuungen der Leistungen der zum Vorphysikum gelangten
 Humanmediziner (36 Pbn)

Abb. 20

A b i t u r n o t e n				K l a u s u r e n		
Deutsch	Mathem.	Gem.K.	Physik.	Chemie	Botanik	Physik
3,36	3,38	3,09	3,13	2,84	9,13	53,18
					83,62	11,11
						53,18
0,68	0,78	0,70	0,76	0,95	2,87	13,75
					41,05	13,75

Mittel-
werte

Streu-
ungen

V o r p h y s i k u m				I-S-T-AMTHAUER	
Physik	Chemie	Zoologie	Botanik	Gesamtergebnis	
2,36	2,62	2,07	2,47	113,5	
56,2	34,9	48,8	36,2		
0,57	0,81	0,72	1,01	3,5	

Mittel-
werte

Streu-
ungen

11) Mittelwerte und Streuungen der Abiturnoten aller 140 Bewerber

Abb. 21

	MW	Streuung
Deutsch	3,49	0,46
Mathem.	3,54	0,80
Gem.K.	3,13	0,78
Physik	3,18	0,87
Chemie	3,10	1,06
Sport	2,15	1,38
Engl.	3,35	1,94
Biologie	2,23	2,02

12) Interkorrelationsmuster der Abiturzeugnisnoten (140 Pbn)

Abb. 22

	Deutsch	Mathem.	Gem.K.	Physik	Chemie	Sport	Englisch	Biologie
D.	1,0000	0,2158	0,4159	0,2359	0,2638			0,3036
M.		1,0000	0,2623	0,2531	0,2405			
G.K.			1,0000	0,4274	0,2196			
Phy.				1,0000	0,3188			
Ch.					1,0000			
S.						1,0000		
E.							1,0000	
B.								1,0000

Signifikant nach t sind Korrelationen von $r \geq 0,1666$.

13) Ergebnisse des I-S-T-AMTHAUER

Abb. 23

		Mittelwert	Streuung
	SE	108,4	2,9
	WA	111,2	8,4
	AN	110,9	6,4
	GE	109,0	5,5
	ME	105,7	10,1
	RA	106,0	9,3
	ZR	106,1	6,0
	FA	108,9	9,3
	WÜ	106,2	5,4
Test gesamt		111,6	4,4
IQ		118	

^{a)} Mittelwert und Streuungen der Items aller Bewerber.

Vergleich der I-S-T-AMTHAUER-Ergebnisse von Pbn mit über-
durchschnittlichen Klausurgesamtergebnissen (35 Pbn) und unter-
durchschnittlichen Klausurgesamtergebnissen (30 Pbn).

Abb. 24

	ÜBERDURCH- SCHNITTLICHE Pbn		UNTERDURCH- SCHNITTLICHE Pbn		Signifikanz der MW-Unterschiede (10 % Niveau)
	Mittelw.	Streuung	Mittelw.	Streuung	
SE	109,0	3,9	107,6	1,8	nicht signifikant
WA	112,6	7,1	109,5	9,7	nicht signifikant
AN	111,8	6,1	109,8	6,8	nicht signifikant
GE	110,3	4,7	107,5	6,0	nicht signifikant
ME	108,2	6,5	102,7	13,1	nicht signifikant
RA	108,0	11,0	103,8	6,7	nicht signifikant
ZR	107,8	4,9	104,2	6,7	signifikant 5 % Ni
FA	111,9	7,2	105,4	9,6	signifikant 10% Ni
WÜ	105,8	5,5	106,6	5,4	nicht signifikant
I-S-T gesamt	113,5	3,5	109,3	4,7	nicht signifikant
IQ	120		114		

- b) Mittelwert und Streuungen der Bewerber getrennt nach Klausur-
leistungen.

) Vergleich der I-S-T-AMTHAUER-Werte für Pbn mit über- und unterdurchschnittlichen Leistungen im Vorphysikum.

Abb. 25

Mittelwerte der

überdurchschnittlichen Pbn		unterdurchschnittlichen Pbn
SE	109,1	107,2
WA	113,6	109,7
AN	111,3	112,0
GE	111,9	106,5 *)
ME	108,1	107,0
RA	108,4	104,0
ZR	107,8	103,6 *)
FA	111,1	109,6
WÜ	106,5	107,5
I-S-T-Gesamt	113,9	111,0

*) Signifikant 10 %-Niveau

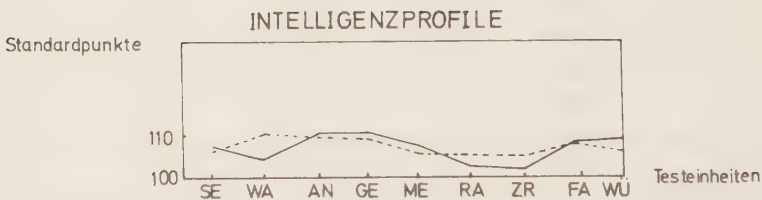
MW des Vorphysikums 2,41

Die Unterschiede der MW der I-S-T-Items wurden nach dem U-Test von MANN und WITNEY auf Signifikanz überprüft. Es ergaben sich auf dem 5 %-Niveau für die Signifikanzschranken 17 und 27 nach Tafeln von SACHS nur für die Items GE mit einem U-Wert von 134,5 und das Item ZR mit einem U-Wert von 147 Signifikanz.

Darstellung der Intelligenzprofile

- d) Darstellung des Intelligenzprofils der untersuchten Probandengruppe anhand der berechneten SW-Punkte in den einzelnen Testitems und Vergleich dieses Profilverlaufes mit dem von AMTHAUER an 36 Ärzten ermittelten.

Abb. 26



IQ der Ärzte 112

IQ der Studenten 118

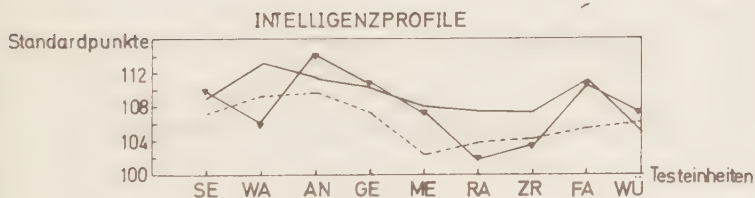
Vergleich der Intelligenzprofile. Durchgezogene Linie: Begabungsprofil von Ärzten nach AMTHAUER; unterbrochene Linie: Begabungsprofil unserer Pbn. Auf der Ordinate AMTHAUER-Standardpunkte; auf der Abszisse die Subtests verschiedener Intelligenzmodalitäten.

SE = Satzergänzungen; WA = Wortauswahl; AN = Analogien;
GE = Gemeinsamkeiten; RA = Rechenaufgaben; ZR = Zahlenreihen;
FA = Figurenauswahl; WÜ = Würfelaufgaben.

Abb. 27

e) Intelligenzprofile als Ergebnisse des I-S-T-AMTHAUER getrennt nach Klausurergebnissen

Ordinate und Abszisse siehe Abb. 26.

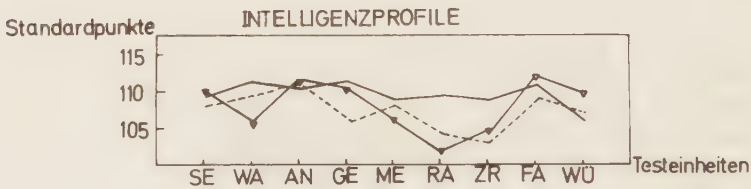


- Profil a: Pbn mit überdurchschnittlichen Klausurergebnissen
IQ - 120
- Profil b: Pbn mit unterdurchschnittlichen Klausurergebnissen
IQ - 114
- ▲ Profil c: "Ärzteprofil" nach AMTHAUER
IQ - 112

Abb. 28

f) Intelligenzprofile - eigene Ergebnisse getrennt nach Vorphysikumsleistungen und Ergebnisse von AMTHAUER

Ordinate und Abszisse siehe Abb. 26.



- Profil a: Pbn mit überdurchschnittlichen Vorphysikumsergebnissen
----- Profil b: Pbn mit unterdurchschnittlichen Vorphysikumsergebnisse
◆—◆— Profil c: "Ärzteprofil" nach AMTHAUER

g) Darstellung der internen HKA des I-S-T-AMTHAUER

Es wurden 4 Faktoren extrahiert und nach Varimax rotiert.

Abb. 29

	1	2	3	4	5
Eigenwerte	3,00	1,15	1,10	0,90	0,83
erklärte Varianz (%)	33 %	13 %	12 %	10 %	9 %

Extraktionskriterium mindestens 70 % aufgeklärte Varianz

	F 1	F 2	F 3	F 4
SE				0,80
WA				0,72
AN	0,74			
GE				0,64
ME			0,91	
RA	0,76			
ZR	0,67			
FA			0,55	0,49
WÜ		0,93		

(eingetragen wurden nur Ladungen von $r \geq \pm 0,5$)

Eine mit denselben Daten berechnete Faktorenanalyse, bei der als Kommunalitätenschätzung der höchste Korrelationskoeffizient verwandt wurde, ergab das gleiche Faktorenmuster. Die Ladungen waren lediglich etwas niedriger. Da durch 4 extrahierte und nach Varimax rotierte Faktoren hier nur etwa 40 % der totalen Varianz erklärt wurden, gelangt oben das Muster der Hauptkomponentenanalyse zur Darstellung.

Abb. 30

14. Korrelationen der 4 Leistungskomplexe

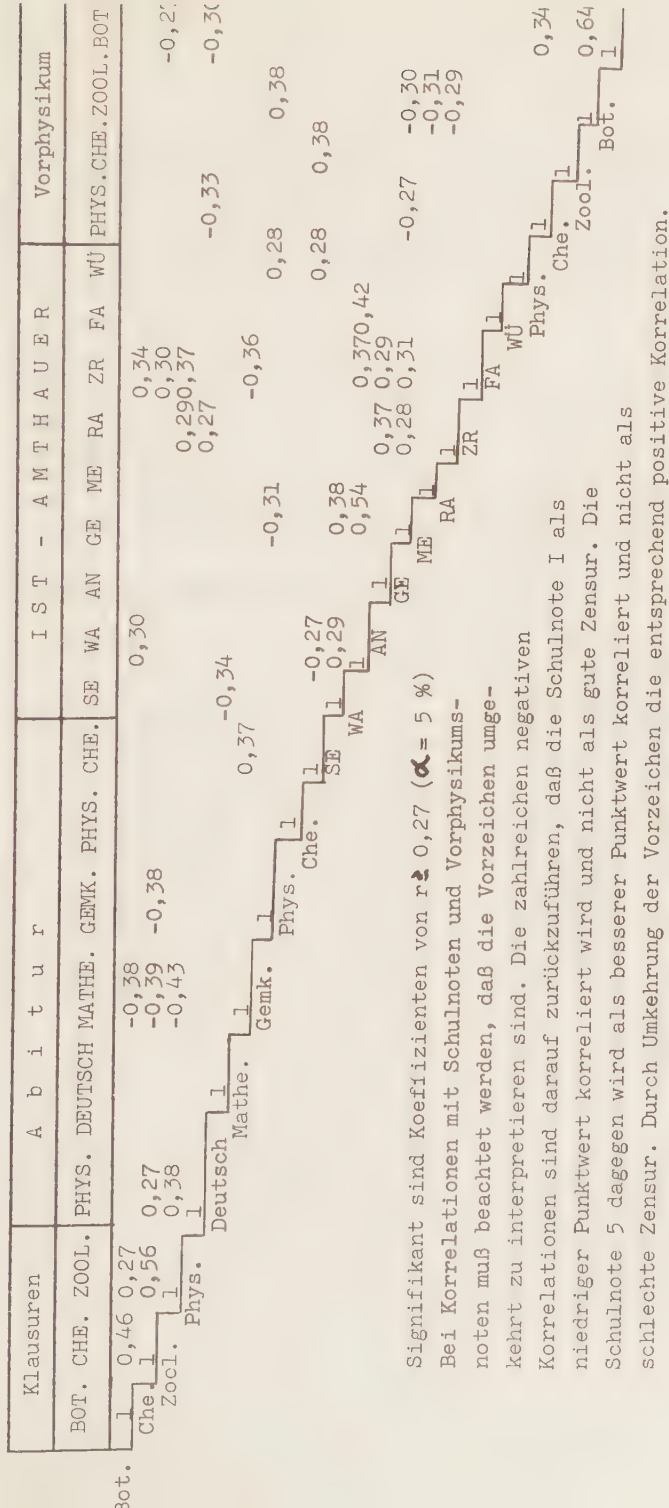
Abitur (MP)	Klausuren **)	Intelligenz	Vorphysikum*)
1,000	0,312	0,148	-0,074
Kl.	1,00	0,315	-0,288
	IST-Amth.	1,000	-0,230
		Vorphy.	1,000

Signifikant bei $r \geq 0,27$

*) Vorphysikum Gesamtnote

**) Ein Klausurgesamtwert wurde berechnet, indem jeder Klausurwert transformiert wurde auf den Mittelwert 0 und die Standardabweichung 1 (Z-Wert). Anschließend wurden die Z-Werte addiert und der Summenscore wieder Z-transformiert.

I) Die Korrelationsmatrix, die aus den Daten eines Pbn-Kollektivs von 36 Studenten berechnet wurde.
Es wurden 22 Variable in die Untersuchung einbezogen.



Ladungsmuster einer Hauptkomponenten-Analyse mit 6 Variablen und 3 Extrahierten und nach Varimax rotierten Faktoren ("Leistungskontrolle").

Abb. 32

Extraktionskriterium mindestens 70 % angeklärte Varianz

	F 1	F 2	F 3
Eigenwerte	3,04	0,82	0,73
erklärte Varianz (%)	50 %	64 %	77 %

	F 1	F 2	F 3	
I-S-T-AMTH.	-	0,90	-	IQ
Botanik	-	0,60	-	K l a u s u r
Chemie	0,74	-	-	
Zoologie	0,87	-	-	
Physik	0,73	-	-	
MP	-	-	0,95	Abi- tur

Eine mit denselben Daten durchgeführte Faktorenanalyse mit dem höchsten Korrelationskoeffizienten als Kommunalitäten-schätzung ergab ein gleiches Faktorenmuster. Die Ladungen waren etwas niedriger, und 3 Faktoren erklärten 52 % der totalen Varianz.

Abb. 33

Faktoren	1	2	3	4	5	6	7	8	9	10
Eigenwert	4,06	2,15	2,02	1,70	1,63	1,39	1,34	1,13	0,96	

Eigenwerte der ersten 9 Faktoren.

Aufgeklärte Varianz beim 9. Eigenvektor = 70 %.

Extraktionskriterium: mindestens 70 %aufgeklärte Varianz.

(Vorzeichen von Noten, z.B. Abitur u. Vorphysikum, sind umgekehrt zu interpretieren).

Das Ergebnis einer Hauptkomponentenanalyse mit 22 Variablen und nach Varimax rotierten Faktoren.

Faktorenmuster mit 9 extrahierten Faktoren

	F 1	F 2	F 3	F 4	F 5	F 6	F 7	F 8	F 9
Botanikklausur	.67	.15	.14	.07	.01	.31	.07	.01	.08
Chemieklausur	.77	.09	.02	.04	.13	-.14	-.03	.16	-.11
Zoologieklausur	.75	.04	.06	.14	.17	.36	-.28	-.00	-.10
Physikklausur	.27	.10	.05	.23	.74	.13	-.08	-.04	-.04
Abitur Deutsch	-.05	.16	-.08	.08	-.01	.82	.21	.02	.05
Abitur Mathem.	-.67	-.10	-.09	.06	-.03	-.10	-.44	-.02	.14
Abitur Gem.K.	-.14	-.37	-.03	-.30	-.20	+.52	-.32	-.26	.15
Abitur Physik	-.04	.13	-.10	.03	.08	-.13	-.82	.08	.08
Abitur Chemie	-.15	-.00	-.38	.23	.05	-.21	-.12	.03	.75
IST-AMTH. SE	.18	.25	.24	.21	-.36	-.45	-.25	-.19	.09
IST-AMTH. WA	.17	.11	.80	.05	-.02	-.08	.09	-.09	-.04
IST-AMTH. AN	-.05	-.09	.05	.85	-.06	-.09	-.03	.08	-.15
IST-AMTH. GE	.05	.31	.56	.24	.10	-.34	.16	-.40	.25
IST-AMTH. ME	.13	.20	.07	.09	-.09	.02	-.05	.86	.04
IST-AMTH. RA	.20	.21	.10	.71	.10	.10	-.02	.03	.22
IST-AMTH. ZR	.55	-.09	.32	.32	-.10	-.08	.06	-.18	-.04
IST-AMTH. FA	.06	-.14	.71	.06	.19	-.01	-.02	.43	-.04
IST-AMTH. WÜ	-.10	-.19	-.26	-.21	.02	.27	-.06	.06	.76
Vorphys.Physik	-.07	-.26	-.23	.32	-.69	.16	.04	.09	-.09
Vorphys.Chemie	-.20	-.17	-.29	.07	.52	-.12	-.01	.50	.23
Vorphys.Zoolog.	-.03	-.92	-.14	-.09	.03	-.03	-.00	-.12	.01
Vprphys.Botanik	-.20	-.79	.09	.07	-.24	-.01	.17	.01	.15

Faktorenmuster der HKA mit 22 Variablen

Abb. 33

	F 1	F 2	F 3	F 4	F 5	F 6	F 7	F 8	F 9
Botanikklausur	x								
Chemieklausur	x								
Zoologieklausur	x								
Physikklausur					x				
Abitur Deutsch						x			
Abitur Mathem.	x								
Abitur Gem.K.						x			
Abitur Physik							-x		
Abitur Chemie									x
IST-AMTHAUER SE									
IST-AMTHAUER WA			x						
IST-AMTHAUER AN				x					
IST-AMTHAUER GE			x						
IST-AMTHAUER ME								x	
IST-AMTHAUER RA				x					
IST-AMTHAUER ZR	x								
IST-AMTHAUER FA			x						
IST-AMTHAUER WÜ									x
Vorphys. Physik					-x				
Vorphys. Chemie					x			x	
Vorphys. Zoologie		-x							
Vorphys. Botanik		-x							

Berücksichtigt: Ladungen 0,5

Ergebnis einer Hauptkomponentenanalyse mit 13 Variablen und 5 Extrahierten und nach Varimax rotierten Faktoren

	1	2	3	4	5	6
Eigenwerte:	3,16	1,88	1,36	1,26	1,09	0,98

Faktorenmuster (berücksichtigt: Ladungen $\geq 0,4$)

Abb. 34

	F1	F 2	F 3	F 4	F 5	
Physik	-	-	-	(-.46)	-	Vor- physikum
Chemie	-	-	-	.82	-	
Zoolog.	-	-	-.93	-	-	
Botan.	-	-	-.79	-	-	
Botan.	.47	-.58	-	-	-	Klau- suren
Chemie	-	-.72	-	-	-	
Zoolog.	-	-.80	-	-	-	
Physik	-	(-.48)	-	.63	-	
Deutsch	-.81	-	-	-	-	Abitur
Mathem.	-	-	-	.51	-	
Gem.K.	-.59	-	(-.37)	-	(.35)	
Physik	-	-	-	-	.85	
Chemie	-	(.38)	-	(.36)	.51	

3 Leistungskomplexe: Vorphysikum - Klausuren - Abitur

VII. Darstellung der Ergebnisse

1. Leistungskomplex Klausuren

a) Ergebnisse der statistischen Testanalyse der Botanikklausur

Die mit Computer ermittelten Teststatistiken ergaben, daß bei 20 Maximal-Punkten ein Mittelwert von 8,3 mit einer Streuung von $\sigma = 2,68$ Punkten erreicht wurde. Der Reliabilitätskoeffizient beträgt $r = 0,45$ und der als Validierung des Testes mit dem Ergebnis der Vorphysikumsprüfung in diesem Fach berechnete Validitätskoeffizient $r = 0,24$ (Abbildung 1), der nicht signifikant ist. Die Verteilung der Rohwerte (Abbildung 2) wird durch eine zweigipfelige Kurve repräsentiert mit einem Einschnitt im Mittelbereich. Sie ist jedoch nicht signifikant verschieden von der Normalverteilung. Die Schwierigkeitsindices (SI) reichten von 4,76 bis 82,1 mit einem mittleren Wert von 36,2 (Abbildung 1). Die Verteilung der SI (Abbildung 6) gibt an, daß 7 von 20 Aufgaben im Schwierigkeitsbereich von 41,0 bis 60,9 liegen. 10 Items befinden sich im Schwierigkeitsbereich darunter, davon 4 im äußerst niedrigen Bereich von 0,0 bis 20,9. Nur drei Aufgaben gehen über den angegebenen Mittelbereich hinaus. Der niedrigste Trennschärfeindex (TI) der Items war 0,08, der höchste 0,47 (Abbildung 1). Die Beziehungen der einzelnen Testitems untereinander kommen im Interkorrelationsmuster der Klausur zum Ausdruck (Abbildung 10). Es zeigt sich, daß Aufgabe 1 und 20 etwas höher korrelieren. Darüber hinaus bestehen nur wenig signifikante Interkorrelationen, ein Zeichen der Heterogenität dieses Testes. Letzteres läßt sich außerdem aus der internen Hauptkomponentenanalyse der Klausur ersehen (Abbildung 14). So erklären z.B. die ersten 3 Faktoren nur 32 % der Varianz. Heterogenität und geringe Zuverlässigkeit stehen ebenfalls in Zusammenhang (LIENERT 1968).

Das Ergebnis der HKA zeigt, daß die substantiellen Ladungen sich auf die 3 Faktoren verteilen, wobei 2 Aufgaben in 2 der Faktoren vertreten sind. Die Items 2 u. 3, wie auch 16, 17, 18, bei denen aus dem Interkorrelationsmuster (Abbildung 12) schon hervorgeht, daß sie keine Beziehungen zu anderen Items

haben, fallen auch im Faktorenmuster auf, da sie mit ihren Ladungen in keinem der erhaltenen Faktoren vorkommen.

b) Statistische Angaben zum Leistungstest in Zoologie

Bei dieser Klausur wurden bei 20 erreichbaren Punkten im Mittel 9,98 erzielt mit einer Streuung $\sigma = 3,1$ Punkten. Die Rohwertverteilung (Abbildung 3) erscheint mehrgipflig, mit einem Tiefpunkt im Mittelbereich und Maxima unter und über dem mittleren Punktwert. Die Normalverteilungsprüfung ergibt jedoch Normalität. Der Koeffizient der Reliabilität beträgt $r = 0,57$ und jener der Validität $r = 0,12$. Letzterer ist nicht signifikant.

Der mittlere SI beträgt 48,87, wobei die Extremwerte $SI = 7,14$ und $SI = 48,5$ sind. Die Häufigkeitsverteilung der SI läßt erkennen (Abbildung 7), daß ein Gipfel im mittleren Schwierigkeitsbereich liegt. 5 Aufgaben liegen im Schwierigkeitsbereich darunter und 8 weit darüber. Die Trennschärfeindizes reichen von 0,7 bis 0,53 (Abbildung 1). Davon liegen 4 Testitems mit ihrem TI im Mittelbereich von 0,40 bis 0,60. Die Restlichen haben niedrigere Trennschärfeindizes (Abbildung 9).

Im Interkorrelationsmuster dieser Klausur fällt auf (Abbildung 11), daß besonders die Aufgaben 2, 9 und 13 mit jeweils 4 signifikanten Korrelationskoeffizienten (0,5 % Niveau) mit 4 anderen Aufgaben in Zusammenhang stehen. Diese Aufgaben sind unter anderem in der internen Hauptkomponentenanalyse (Abbildung 15) im ersten Faktor vertreten. Die Items Nr. 5, 11, 18, 19, die im Interkorrelationsmuster (0,1 % -Niveau) keinerlei Korrelationen zu anderen Items haben, fehlen dagegen in den extrahierten Faktoren der internen HKA.

c) Statistische Ergebnisse der Physiklausur

In diesem freibeantwortbaren Wissenstest wurden von 100 erreichbaren Punkten im Mittel 49 Punkte bei einer Streuung von $\sigma = 21$ erzielt (Abbildung 1). Ein Reliabilitätskoeffizient wurde als r_{tt} berechnet. Die Korrelation mit der Physikprüfung des Vorphysikums als Validitätskriterium beträgt $r = 0,33$. Sie ist nicht signifikant. Für den Reliabilitätskoeffizienten wurde nach der Split-half-Methode ein Wert von 0,999 ermittelt.

Die Verteilung der Rohwerte (Abbildung 4) zeigt, daß fast die Hälfte der Probanden mittlere Punktwerte im Bereich von 40 bis 60 Punkten erreichten. Etwa 32 % der Bewerber liegen darunter, 20 % darüber. Die Normalverteilung konnte auch hier statistisch gesichert werden. Das Schwierigkeitsintervall der Aufgaben geht von $SI = 29$ bis $SI = 72,3$ mit einem mittleren SI von $56,23$. Als niedrigster TI kommt ein Wert von $0,06$, als höchster einer von $0,76$ (Abbildung 1) vor. Nur 2 Aufgaben liegen im unteren Schwierigkeitsbereich von SI $21,0$ bis $40,9$ (Tabelle 8). Die TI gruppieren sich (Tabelle 9) hauptsächlich, d.h. zu 70 % um und über den Mittelbereich. Aus dem Interkorrelationsmuster läßt sich entnehmen (Abbildung 13), daß besonders die Subtests 1, 9 und 10 sehr zahlreich mit anderen in Beziehung stehen, und zum Teil mit Werten um $r = 0,4$ korrelieren. Als gänzlich unabhängig erweist sich nur Subtest 2. Seine Besonderheit kommt auch in der internen HKA des Testes zum Ausdruck (Abbildung 16). Hier zeigt er im 3. Faktor eine substantielle Ladung von $-0,73$. Während die Aufgabengruppen 1 und 10 im ersten Faktor, zusammen mit Aufgabe 4, sehr hoch mit Ladungen von $0,76$, $0,72$, $0,78$ (Abbildung 16) vertreten sind, findet sich der 9. Aufgabenkomplex im zweiten Faktor mit einer Ladung von $0,55$ wieder, zusammen mit dem Subtest 3 mit $0,83$ und 7 mit $0,67$.

d) Ergebnis der statistischen Untersuchung der Chemieklausur

In 10 Aufgaben waren 200 Punkte erreichbar. Der niedrigste erzielte Rohwert war 4 Punkte, der höchste 190 Punkte. Der Mittelwert beträgt $69,2$ mit einer Streuung von $\sigma = 18,8$ (Abbildung 1). Die Rohwerteverteilung läßt erkennen (Abbildung 4), daß mit $18,9$ % aller Bewerber das Maximum der erreichbaren Punkte im Bereich von 20 bis 40 liegt. Mittlere Punktwerte von 90 und 110 finden sich nur bei $11,5$ % der Klausurteilnehmer. $56,5$ % der Bewerber liegen unterhalb dieses Mittelwertes. Ergebnisse von 140 bis 190 Punkten erreichten $3,5$ % der Studenten. Die Normalverteilungsprüfung ergab, daß die erreichten Punktzahlen sich nicht normal verteilen. Auch bei dieser Klausur ist der Validitätskoeffizient auf dem 5 % - Niveau nicht signifikant. Der Reliabilitätskoeffizient ergab nach der Split-half-Methode einen Wert von $0,998$.

Die Aufstellung der SI (Tabelle 8) läßt erkennen, daß die SI in Bereichen von 10 bis 50 liegen. Kein Wert geht darüber hinaus. Bei den TI (Tabelle 9) findet sich dagegen kein Index unter 0,4. 80 % aller TI's liegen im Intervall von 0,50 bis 0,80. Das Interkorrelationsmuster (Abb. 12) hat als Ergebnis, daß fast ausnahmslos jede Aufgabe mit jeder anderen signifikant und zum Teil hoch korreliert. Die interne HKA (Abb. 17) zeigt bei 3 rotierten Faktoren, daß jede Aufgabe an der Varianzerklärung einer der drei Faktoren beteiligt ist. Die Items 5 und 7 laden den ersten Faktor mit 0,78 und 0,82 besonders hoch. Im zweiten Faktor ist die Variable (8 $r = 0,88$) sehr stark vertreten und im dritten die Items 4 mit 0,73 und 9 mit 0,80.

2.) Leistungskomplex: Abitur

Wie aus Abbildung 21 zu erkennen ist, zeigen die Mittelwerte der Abiturnoten aller 140 Bewerber an, daß die Zensuren der sprachlichen Hauptfächer wie Deutsch, Englisch, aber auch die naturwissenschaftlichen Fächer Mathematik, Physik, Chemie zwischen 3 und 4 liegen. Die Note in Englisch streut mit $\sigma = 1,94$ besonders stark. In Sport und Biologie sind die MW etwas unter 2, streuen jedoch ebenfalls $\sigma = 1,38$ für Sport und $\sigma = 2,02$ für Biologie sehr stark. Die Prüfung auf Normalverteilung der Noten ergab Folgendes: Die MP sind auf dem 10 % - Signifikanzniveau nicht signifikant von der Normalverteilung verschieden. Ebenfalls normalverteilt sind die Noten in Gemeinschaftskunde, Mathematik und Physik. Die Chemiezensur ist nur auf dem 1 % - Signifikanzniveau normalverteilt, auf höherem Niveau nicht mehr. Nicht normalverteilt ist bemerkenswerterweise die Abiturnote in Deutsch. Die beste Zensur wurde im Fach Chemie mit 2,84 erreicht, die schlechteste in Mathematik mit 3,36.

Das Interkorrelationsmuster der Abiturnoten aller Bewerber (Abb. 22) zeigt im besonderen, daß Sport und Englisch keine Korrelationen mit anderen Abiturfächern haben. Die Note in Deutsch dagegen korreliert mit Gemeinschaftskunde zu $r = 0,416$ und Biologie zu $r = 0,304$. Die Physiknote korreliert

mit Gemeinschaftskunde zu $r = 0,427$, aber auch mit der Leistung in Chemie zu $r = 0,319$. Wie aus Abbildung 27 hervorgeht, korreliert die Mathematiknote mit Physik ($r = 0,25$).

3. Leistungskomplex I-S-T-AMTHAUER

MW und Streuung der einzelnen I-S-T-Subtests aller Probanden, die sich diesem Test unterzogen, sind aus Tabelle 23 ersichtlich. Der höchste Punktwert von 111,2 wurde dabei in "WA" erzielt, der niedrigste mit 105,7 in "ME". Der mittlere Standardpunktwert des Gesamttestes beträgt 111,6 mit einer Streuung von $\sigma = 4,4$. Es entspricht dieses einem IQ von 118.

Die Prüfung auf Normalverteilung der von allen Pbn erreichten Punktwerte in den einzelnen Items ergab:

- a) auf dem 5 % - Signifikanzniveau sind normalverteilt die Subtests: WA; GE; ME; RA; ZR; FA; WÜ; und der durch Transformation der Gesamtstandardwerte berechnete IQ. Nicht normalverteilt sind auf dem 5 % - Niveau: SE und AN.
- b) auf dem 10 % - Niveau sind nicht normalverteilt die Items: SE; AN; GE; ME; FA.

Da verschiedene Werte sich somit als nicht normalverteilt herausstellten, wurde zur Prüfung der Signifikanz der MW-Unterschiede der parameterfreie Test nach MANN und WITNEY (SACHS) durchgeführt.

Vergleicht man nun die I-S-T-AMTHAUER-Werte der Pbn mit über- und unterdurchschnittlichen Klausurgesamtergebnissen (Abb. 24), so ergibt sich ein signifikanter Leistungsunterschied beider Gruppen auf dem 10 % - Signifikanzniveau nur für die Items FA und ZR. Letzteres ist auch auf dem 5 % - Niveau signifikant, allerdings nur schwach.

In einer weiteren Untersuchung wurde das Probandenkollektiv der Bewerber, die aufgenommen wurden und die Vorphysikumsprüfung ablegten, geteilt nach über- und unterdurchschnittlichen Kandidaten und die jeweiligen I-S-T-Ergebnisse verglichen (Abb. 25). Es zeigt sich, daß die Studenten mit besseren Vorphysikumszensuren auch etwas höhere AMTHAUER-Ergebnisse haben. Diese Unterschiede konnten allerdings nur für die Items GE und ZR als statistisch signifikant abgesichert werden.

Im Interkorrelationsmuster des Testes (Abb. 19), das am Ausgangskollektiv ermittelt wurde, fällt auf, daß die Items "ME" und "WÜ" nicht mit anderen korrelieren. Die restlichen 7 Items korrelieren untereinander mehr oder weniger.

Das interne Faktorenmuster des AMTHAUER (Abb. 29) sowohl als Hauptkomponentenanalyse mit 5 extrahierten und nach Varimax rotierten Faktoren, wie auch als Ergebnis einer FA mit dem höchsten Korrelationskoeffizienten als Kommunalitätenschätzung zeigt (Abb. 32), daß im ersten Faktor die Items "AN", "RA" und "ZR" mit Ladungen von 0,65 und 0,75 vertreten sind. Der zweite Faktor enthält nur eine Ladung, und zwar "WÜ" mit 0,93. Die Varianz des dritten Faktors wird im wesentlichen von den Items "ME" (0,91) und "FA" (0,55) erklärt. Den vierten Faktor laden hauptsächlich 4 Testteile, darunter besonders hoch "SE" mit 0,80 und "WA" mit 0,72.

4. Leistungskomplex: Vorphysikum

Aus den einzelnen Prüfungen des Vorphysikums liegen Mittelwerte und Streuung als Ergebnisse vor (Abb. 20). Die mündliche Prüfung in Zoologie erbrachte die beste Durchschnittszensur mit 2,07. Die Botanikprüfung mit einem MW von 2,47 streut mit $\sigma = 1,01$ am stärksten. Die Prüfung in Chemie erbrachte die niedrigsten Zensuren, deren MW 2,62 beträgt. In der mündlichen Prüfung in Physik wurde eine durchschnittliche Zensur von 2,36 erreicht bei einer Streuung von $\sigma = 0,57$.

5. Statistische Analyse der Wechselbeziehungen des gesamten Variablenkollektivs

Über einige interessante Beziehungen einzelner Testitems der Klausuren zu den übrigen Variablen wird aus Gründen der Übersichtlichkeit nur im Zusammenhang mit der unmittelbaren Diskussion der jeweiligen Items berichtet.

Die Analyse der Zusammenhänge zwischen Abiturkomplex und Klausurkomplex (Abb. 18) zeigt, daß alle 4 naturwissenschaftlichen Leistungsprüfungen untereinander und mit den Marburger Punkten signifikant korrelieren. Darüber hinaus fällt auf, daß die Abiturnote in Mathematik signifikant positiv mit

den Klausurleistungen korreliert, dabei am höchsten mit der Chemieklausur ($r = -0,45$) und am niedrigsten mit der Physikklausur ($r = -0,26$). Außerdem korrelieren alle anderen Klausuren mit dem Gesamtergebnis des I-S-T-AMTHAUER (Abb. 19), also mit dem IQ. Die Zoologie ist dabei am geringsten repräsentiert ($r = 0,26$), der Leistungstest in Botanik dagegen am meisten ($r = 0,42$). Darunter zeichnet sich besonders die Beziehung zwischen Botanik und folgenden Items auf: "GE" ($r = 0,38$), "ZR" ($r = 0,36$), "FA" ($r = 0,31$). Mit der Merkfähigkeit ("ME") korrelieren besonders die Chemieklausur ($r = 0,33$) und die Zoologieklausur ($r = 0,27$).

Über die Bedeutung der Wechselbeziehungen aller 4 Leistungskomplexe gibt Tabelle 31 Auskunft. Hier wurden die Klausurergebnisse, die Abiturnoten, der Intelligenz-Struktur-Test, und die Vorphysikumsnoten korreliert. Das Datenmaterial bezieht sich daher nur auf das bis zum Vorphysikum und weiter gelangte Studentenkollektiv von 36 Probanden. Außer den oben bereits erwähnten Korrelationen der Klausuren mit dem Abitur und dem I-S-T-AMTHAUER ergibt sich hier, daß die Physikklausur Bedeutung für das Vorphysikum hat. So korrelieren beide Prüfungen miteinander ($r = 0,33$) signifikant. Die naturwissenschaftliche Vorprüfung in Chemie dagegen korreliert mit der Abiturnote in Chemie ($r = 0,38$). Die Vorphysikumsprüfung in Zoologie korreliert mit dem Abiturfach Gemeinschaftskunde zu $r = 0,38$.

Betrachtet man abschließend die 4 in der Analyse vertretenen Leistungsgruppen als Gesamtkomplexe und ihre Beziehungen untereinander (Abb. 30), so ergab sich für die Daten unseres Kollektivs, daß Abitur und schriftliche naturwissenschaftliche Leistungskontrollen mit $r = 0,312$ nur signifikant auf dem 0,1 % - Signifikanzniveau korrelieren. Das Abitur hat zur Intelligenz und zum Vorphysikum hier keine Beziehung. Die Klausuren haben außer obiger Beziehung eine ebenso schwach signifikante Korrelation zur Intelligenz. Der Leistungskomplex Vorphysikum hat hier keine signifikanten Korrelationen mehr zu den anderen Testgruppen.

Um die Struktur des Variablenkollektivs zu erhellen, wurden Faktorenanalysen berechnet, und zwar nach der Hauptachsenmethode, dann mit Kommunalitätenschätzungen mittels des höchsten Korrelationskoeffizienten oder des multiplen Korrelationskoeffizienten. Es zeigt sich, daß die Ladungsmuster der Faktoren in etwa gleich waren. Für die Darstellung und Interpretation wurden Übersichtlichkeit wegen nur einige Ergebnisse ausgewählt.

Als grobe Übersicht wurden zunächst die 3 großen Leistungskomplexe, denen ein großes Pbn-Kollektiv zugrunde liegt, und ihre Verteilung auf die Faktoren untersucht. Es sind dies: das Abitur (MP), die 4 Klausuren und der I-S-T-AMTHAUER.

Die Hauptkomponentenanalyse mit 6 Variablen und 3 extrahierten und nach Varimax rotierten Faktoren (Abb. 32) zeigt, daß der 1. Faktor hohe Ladungen der Klausuren enthält, und zwar von Chemie mit 0,74, von Zoologie mit 0,87 und von Physik mit 0,73. Im 2. Faktor laden besonders der I-S-T-AMTHAUER mit 0,90 und die Botanikklausur, die in der 1. Dimension nicht enthalten ist, mit 0,60. Der 3. Faktor wird mit 0,95 durch die MP repräsentiert.

Erweitert man die obige Untersuchung durch Einbeziehung der 9 I-S-T-Items und der 4 Vorphysikumsnoten und rotiert 9 Faktoren, so ergibt sich folgendes Faktorenmuster (Abb. 33). Die Varianz des 1. Faktors wird im wesentlichen durch die Klausuren erklärt, und zwar die Botanikklausur 0,67, die Chemieklausur 0,77, die Zoologieklausur 0,75, außerdem haben die Abiturnote Mathematik mit 0,67 und der Zahlenreihentest des I-S-T- mit 0,55 daran teil.

Im 2. Faktor laden nur Vorphysikumsprüfungen, die Zoologie mit 0,92 und die Botanik mit 0,79. Im 3. Faktor findet man nur I-S-T-AMTHAUER-Items; "WA" lädt mit 0,80, "FA" mit 0,71. Ebenso beim 4. Faktor, hier laden "AN" mit 0,85 und "RA" mit 0,71. Die Varianzanteile des 5. Faktors werden durch Physikprüfungen erklärt, so durch die Klausur mit 0,74 und die Vorphysikumsprüfung mit 0,69. Im 6. und 7. Faktor stellen sich besonders Abiturnoten dar, z. B. Deutsch 0,82 und Ge-

meinschaftskunde 0,52 im Sechsten und Physik 0,82 im Siebten. An 8. Stelle steht ein gemischter Faktor mit den Ladungen "ME" 0,86 des Intelligenztestes und der Vorphyskumsprüfung in Chemie 0,50. Im 9. Faktor tritt die Abiturnote in Chemie mit 0,75 zusammen auf mit "WÜ" 0,76 des I-S-T.

VIII. Diskussion der Ergebnisse

1. Klausuren

a) Botanikklausur

Wie aus der Darstellung der Ergebnisse hervorgeht, ist die Botanikklausur mit ihren im Mittel erreichten 36,2 % der Punkte als schwierig anzusehen. Die Ursache könnte darin zu suchen sein, daß einerseits der Bewerberdurchschnitt mit der Beantwortung der zum Teil sehr schweren Biochemiefragen (wie bereits oben angedeutet) überfordert war, andererseits die Testkonstruktion mit einigen schlecht konstruierten Items nicht ganz ohne Wirkung auf das Ergebnis ist.

Die Rohwerteverteilung zeigt für unser Kollektiv zwar, daß gerade im Mittelbereich der erreichbaren Punkte ein Tiefpunkt liegt, d. h. also, daß nur relativ wenige Pbn mittlere Punkte erreichten, doch ergibt die Prüfung auf Normalverteilung der Werte, daß sie nicht signifikant von der Normalverteilung verschieden sind. Die Inzisur ist somit rein zufälliger Natur.

Der Reliabilitätskoeffizient ist - da nach SCHRADER (1968) Koeffizienten unter 0,80 als bedenklich anzusehen sind - sehr unbefriedigend.

Wenn eine Vortestung stattgefunden hätte, wie von UEXKÜLL für Tests mit Selektionsfunktion fordert (v. UEXKÜLL 1968), wäre die Reliabilität des Testes eventuell besser geworden. Ein weiterer Einfluß auf die Reliabilität geht vor allem von den Trennschärfeindices der Aufgaben aus. So vermindern niedrige Trennschärfeindices die Reliabilität der Tests (v. NÄERSSSEN 1968). Auch HUBBARD weist darauf hin, daß die Reliabilität eines Testes mit der wachsenden Zahl schlecht konstruierter Items sinkt. Überdies wurde der Koeffizient nach der Konsistenzmethode von KUDER-RICHARDSON geschätzt (HUBBARD 1968), in den die Homogenität des Testes eingeht. Wie die Item-Interkorrelationsmatrix des Testes zeigt, sprechen die Interkorrelationen für eine Heterogenität dieser Klausur. Darüber hinaus hätte ein größeres Fragenkompodium diesen Koeffi-

zienten sicherlich erhöht (v. UEXKÜLL 1968; INGENKAMP 1968; HUBBARD 1968). Für Leistungstests mit Multiple-Choice-Fragen sind 20 Items ungewöhnlich wenig, so daß die Ergebnisse möglicherweise für ein Individuum nicht interpretiert werden können (INGENKAMP 1968). In Amerika, wo man die bisher längsten Test-erfahrungen hat, bestehen diese Leistungskontrollen zumeist aus weit mehr Fragen. So gibt HUBBARD z.B. als Richtzahlen für eine Abschlußprüfung in einem Fach etwa 160 bis 200 Fragen an (HUBBARD 1968), die in etwa 2 Stunden zu beantworten sind. Bei der Konstruktion dieser Aufgaben beteiligt sich eine erfahrene Expertengruppe aus jeweiligen med. Fachspezialisten, Psychologen und Statistikern (HUBBARD 1968; KAPUSTE 1966; INGENKAMP und MARSOLEK 1968).

Die Validität der Botanikklausur als "kriterienbezogene Validität" sollte am Studienerfolg, bei uns also am Vorphysikum gemessen werden. Der Koeffizient fiel erwartungsgemäß niedrig aus, allerdings so niedrig, daß er nicht einmal signifikant ist. In negativer Weise wirkt sich dabei aus, daß das Vorphysikum eine mündliche Prüfung darstellt, mit allen imponderabilen Nachteilen dieser Prüfungsart (SCHULZ, REMSCHMIDT, PRINZ 1970; STIMPEL 1969). Außerdem kann eine Stichprobengröße von 84 Probanden zu einem relevanten Stichprobenfehler des Koeffizienten führen. Darüber hinaus beeinflußt die Selektion den Validitätskoeffizienten (SCHRADER 1968 b) z.B. dann, wenn der Koeffizient an der Gruppe von Vorphysikumskandidaten gewonnen wird, die zum Teil auf der Grundlage von Werten in gerade diesem Test selektiert wurden, der validiert wird.

Wie beim MW bereits erwähnt, war die Klausur relativ schwer, was sich auch in den Schwierigkeitsgraden der Testitems widerspiegelt. Die Verteilung der Aufgabenschwierigkeiten zeigt ein deutliches Minimum im Bereich mittlerer SI. Normalerweise werden Tests so standardisiert, daß die Verteilung der SI annähernd normal ist, d.h. mittlere Aufgabenschwierigkeiten am häufigsten vertreten sind, da diese Items am besten differenzieren (HUBBARD 1968; SCHRADER 1968). Will man einen Test aus bestimmten Gründen, wie z.B. der Selektion, so gestalten, daß ein bestimmter Prozentsatz von Probanden unterhalb eines bestimmten Testpunkt-

wertes eliminiert werden soll, wählt man die SI so, daß eine möglichst exakte Diskrimination im kritischen Bereich möglich ist. HUBBARD (1968) gibt als Beispiel an, daß ein Test, der die unteren 15 % von Probanden deutlich von den oberen 85 % trennen soll, überwiegend Aufgaben vom SI = 85 haben sollte. Die hier vorliegende Verteilung ist jedoch wahrscheinlich ein Zufallsergebnis, da die Aufgabenschwierigkeiten vorher nicht bekannt waren.

Da bekanntlich eine parabolische Beziehung zwischen SI und TI der Items besteht (LIENERT 1967), aus der zu entnehmen ist, daß sehr schwere und sehr leichte Aufgaben, wie sie bei uns hauptsächlich vorliegen, wenig scharf trennen, ist das Ergebnis der hier erhaltenen TI entsprechend ungünstig. Kein einziges der hier vorhandenen Items erreicht besonders gute Indexbereiche über 0,5. Genau die Hälfte aller Aufgaben haben Trennschärfeindizes unter 0,30 und sind somit "von zweifelhafter Brauchbarkeit" (SCHRADER 1968).

Als Beispiel einer Item-Analyse und Aufgabenbeurteilung sollen zunächst einige Fragen dieses Testes diskutiert werden. Zur Beurteilung einer Aufgabe schlägt HUBBARD vor, bestimmte Fragen an das Item zu stellen. Im folgenden soll eine Botanikaufgabe mit Text und Alternativen sowie auch die Ergebnisse ihrer itemstatistischen Analyse vorgestellt werden:

Aufgabe Nr. 6 in der Klausur

Text:

Welche Gärung verläuft aerob?

- a. Alkoholische Gärung
- b. Milchsäuregärung
- c. Essigsäuregärung
- d. Buttersäuregärung
- e. Keine Antwort ist richtig

Ergebnisse der Itemstatistik:

Nicht beantwortet:	1,19	0,00 (s)	1,19 (g)
Schwierigkeitsindex:	19,05		
Trennschärfeindex:	0,08		

Häufigkeit der Beantwortung	Leistungskorrelation			
Alternative a:	28,57	15,48 (s)	13,10 (g)	- 0,05
Alternative b:	11,09	7,14 (s)	4,76 (g)	- 0,10
*Alternative c:	19,05	10,71 (s)	8,38 (g)	+ 0,08
Alternative d:	4,76	3,57 (s)	1,19 (g)	- 0,00
Alternative e:	34,52	20,24 (s)	14,20 (g)	+ 0,04

s = Pbn mit Testergebnis unter dem MW der Rohwerte

g = Pbn mit Testergebnis über dem MW der Rohwerte

* = Richtige Antwort

HUBBARD fragt unter anderem:

1. Behandelt die Frage einen wichtigen Aspekt eines bestimmten Tatbestandes?

Wenn man auch geneigt ist, auf diese Frage spontan eine positive Antwort zu geben, da es sich eindeutig um ein sehr wichtiges biochemisches Problem handelt, so gilt es doch zu bedenken, daß nur nach einem wichtigen Aspekt gefragt wird. Denn das ist ja ein Charakteristikum des Multiple-Choice-Verfahrens, daß größere Problemkomplexe in viele Einzelfragen aufgelöst werden und jedes Item nur einen Teilaspekt erfaßt. Bei dieser Beispielaufgabe dagegen bezieht sich die Frage auf ein sehr komplexes Geschehen und überfordert so den Gefragten.

2. Prüft die Frage nicht nur unwesentliches lexikalisches Wissen?

Die Frage kann verneint werden, da es sich um wesentliche biologische Vorgänge handelt, die man - um sie überhaupt lernen und reproduzieren zu können - schon weitgehend verstanden haben muß.

3. Ist das zentrale Problem klar und exakt ausgedrückt?

Bezüglich der Formulierung des erfragten Problems ist diese Aufgabenstellung beispielhaft gut.

4. Sind doppelte Verneinungen vermieden?

Ja!

5. Ist die Frage so klar und knapp wie nur möglich formuliert?

Bei dem hiergenannten Beispiel ergibt die Prüfung der Frage die gleiche Antwort wie bei Frage 3 (s.o.).

6. Würde das Item dadurch verbessert, daß die Alternativen "keine Antwort ist richtig" hinzugefügt wurde?

HUBBARD fordert für diese Antwortmöglichkeit, daß sie den Probanden zwingt, über das gesamte Problem, d.h. über die angebotenen Alternativen hinaus noch einmal nachzudenken.

Diese Forderung trifft bezüglich der hier vorliegenden Aufgabe in keiner Weise zu. Wie der Schwierigkeitsindex zeigt, handelt es sich um ein extrem schweres Item, bei der die wenigsten Probanden die Lösung wissen können. Sie können also auch nicht ange-regt werden, außer den 4 angegebenen noch andere Möglichkeiten zu überdenken und zu überprüfen.

7. Ist jede Antwort eine plausible Erklärung?

Gewiß ist jede Antwort eine kritisch zu prüfende Möglichkeit, doch verlangt sie vom Pbn nicht nur das komplexe zentrale Problem zu überdenken, sondern mit jeder Alternative neue, nicht viel weniger schwierige Prozesse zu kennen.

8. Entspricht die Frage dem zu erwartenden Wissensstand eines Prüflings?

Diese Frage ist zu verneinen.

Inhaltlich geht es in dem Item darum, 4 verschiedene sehr spezifische bio-chemische Prozesse differenzieren zu können. Um also die richtige Antwort geben zu können, mußte der Proband über 4 sehr verschiedene Prozesse sehr genau Bescheid wissen. Da dieser Fragentyp jedoch ganz besonders dazu bestimmt ist, die Kenntnisse in bezug auf eine einzige ganz bestimmte Problematik zu überprüfen und nicht 4 verschiedene Themen in einem Item zu erfragen (HUBBARD 1968), waren die Studenten allein schon von der Itemkonstruktion her überfordert. Da außerdem die Mehrzahl der Bewerber nach dem 1. naturwissenschaftlichen Semester an dieser Klausur teilnahm, hatten sie noch keine vorklinische Ausbildung in Biochemie erfahren können. Im Rahmen des Botanikunterrichtes werden die biochemischen Prozesse der pflanzlichen Natur zwar er-

klärt, doch kann ein derart differenziertes Verständnis, wie es in dieser Aufgabe gefordert wurde, wegen der mangelnden Grundlagen nicht generell vorausgesetzt werden. Dementsprechend fällt die Itemstatistik dieser Aufgabe erwartungsgemäß schlecht aus. Als wahrscheinlichste Lösung wurde von etwa 35 % der Studenten, darunter vor allem auch von 15 % der Bewerber mit insgesamt gutem Klausurergebnis die Alternative e gewählt. Die richtige Lösung wurde von 19 % der Probanden gewählt, darunter von 11 % der Pbn mit überdurchschnittlichen Testergebnissen. Wie bereits angedeutet, kann es sich hierbei um reine Zufallstreffer handeln. Als sehr starker Ablenker erweist sich die Alternative a, die von knapp 30 % der Bewerber gewählt wurde, darunter von über 13 % überdurchschnittlichen Testteilnehmern. Dieses Testitem fällt darüber hinaus im Interkorrelationsmuster auf, da keinerlei signifikante Beziehungen zu irgendeinem der übrigen Variablen bestehen.

Im Falle einer Vortestung wäre eine solche Aufgabe im Rahmen einer Leistungstestbatterie mit Selektionsfunktion gewiß rechtzeitig eliminiert worden.

Im Gegensatz zu dem eben diskutierten Item wird im folgenden eine Aufgabe der selben Klausur vorgestellt, die sowohl allen Forderungen HUBBARD's an ein gutes Item genüge zu leisten vermag, außerdem eine zufriedenstellende Itemstatistik erbrachte und sich auszeichnet durch zahlreiche Korrelationen mit anderen in dieser Testbatterie analysierten Variablen.

Aufgabe Nr. 8 in der Klausur

Frage:

In welchen Strukturelementen der Zelle sind die Enzyme des Zitronensäurezyklus lokalisiert?

- a. In den Plastiden
- b. In den Dictyosomen
- c. Im endoplasmatischen Reticulum
- d. In den Mitochondrien
- e. Keine Antwort ist richtig

Nicht beantwortet: 1,19 (s) 0,00 (g)
 Schwierigkeitsindex: 59,52
 Trennschärfeindex: 0,3717

	Häufigkeit der Beantwortung			Leistungskorrelation
Alternative a:	10,71	7,14 (s)	3,57 (g)	-0,0815
Alternative b:	11,90	9,52 (s)	2,38 (g)	-0,1504
Alternative c:	3,57	3,57 (s)	0,00 (g)	-0,0692
* Alternative d:	59,52	26,19 (s)	31,33 (g)	0,3717
Alternative e:	13,10	9,52 (s)	3,57 (g)	-0,279

Untersucht man auch dieses Testitem nach dem HUBBARDSchema (s.o.), so kommt man zu dem Ergebnis, daß es den Anforderungen, die an ein gutes Item gestellt werden, voll Genüge zu leisten vermag. Die Itemstatistik dieser Aufgabe ist entsprechend zufriedenstellend. Darüber hinaus fällt dieses Item durch zahlreiche Korrelationen mit anderen Variablen dieser Testbatterie auf. Zur Chemie z.B. beträgt die Korrelation $r = 0,27$, zur Botanik korreliert dieses Item mit $r = 0,36$ (entspricht dem TI), mit der Klausur in Zoologie korreliert dieses Item zu $r = 0,24$, außerdem mit dem Abitur (MP) zu $r = 0,29$. Eine Korrelation zu $r = 0,27$ mit dem I-S-T-AMTHAUER-Item "ME" deutet seine Abhängigkeit von der Merkfähigkeit an. Die statistischen Kennziffern dieses Items sind fast optimal. Bei einem SI von 59,5 ist der Subtest überdurchschnittlich leicht. Der TI von 0,37 gibt an, daß die Aufgabe ausreichend gut zwischen Probanden mit überdurchschnittlichem und unterdurchschnittlichem Gesamtergebnis trennt. Kein einziger guter Proband hat die Aufgabe nicht beantwortet, und nur die Richtigerantwort hat positive Leistungskorrelation. Die Ablenker sind ebenfalls recht gut. Keiner ist zu sehr oder zu wenig attraktiv. Es findet sich hier also eine Aufgabe, die für verschiedene naturwissenschaftliche Bereiche, darunter auch besonders der Medizin, von großem Interesse ist, gepaart mit einer beispielhaft guten Itemkonstruktion.

Das Interkorrelationsmuster dieser Klausur bringt den Heterogenitätscharakter dieses Testes zum Ausdruck, da 16 Items nur wenige und niedrige signifikante Korrelationen zu anderen Items aufweisen und 4 Aufgaben ohne Beziehungen zum übrigen Test sind. Der nächste von der Inter-Item-Korrelationsmatrix des Testes

ausgehende Schritt zur Analyse der Aufgabenstruktur ist die Faktorenanalyse des Testes (HENRYSSON 1968). Da der hier vorliegende Test eventuell verschiedene inhaltliche Bereiche erfaßt, wurde die Faktorenanalyse berechnet, um zu überprüfen, ob trotz der Heterogenität der Klausur die inhaltliche Dreigliederung auch in dem Muster der Faktorenladungen dreier extrahierter und rotierter Faktoren zum Ausdruck kommt. Der durch diese 3 Faktoren erklärte Varianzanteil ist viel geringer als z.B. bei den beiden Tests mit freier Beantwortung. Da die Variablen an den Ecken eines mehrdimensionalen Würfels gelegen sind, ist es schwer, durch Rotation die Achsen so zu erhalten, daß jeweils ein Maximum an Varianz erklärt wird. Das Ergebnis der internen HKA zeigt, daß der erste Faktor mit Aufgaben geladen ist, die allgemein botanisch-biochemisches Wissen voraussetzen, jedoch ohne spezielle Biochemiekenntnisse gelöst werden können. Die zweite Dimension erfaßt dagegen ganz deutlich Leistungen jener Aufgabengruppe, die zum Teil sehr detaillierte Kenntnisse ganz spezieller biochemischer Prozesse der pflanzlichen Natur voraussetzten. Der dritte Faktor ist dagegen ein ganz allgemeiner botanischer Leistungsfaktor, so daß die von der inhaltlichen Gruppierung der Klausur her erwartete deutliche Dreidimensionalität des Testes wegen der Heterogenität nur angedeutet wird.

b) Itemanalyse der Zoologieklausur

Die statistische Analyse des zweiten Leistungstests hat zum Ergebnis, daß die Zoologieklausur von durchschnittlicher Schwierigkeit war. Dieses spiegelt sich wider im guten MW der erreichten Punkte, sowie in den Aufgabenschwierigkeiten. Der MW der SI liegt genau im mittleren Schwierigkeitsbereich, und die Variationsbreite ist etwas geringer als im Botaniktest. Die Häufigkeitsverteilung der SI ist zwar etwas unregelmäßig, aber doch mit deutlichem Maximum in der Mitte, d.h. bei 50 %. Die Rohwertverteilung sieht ähnlich aus, sie ist normal verteilt. Bei den Trennschärfeindices der Aufgaben fallen nur 3 Items durch mangelnde Diskrimination auf. Die übrigen Aufgaben trennen zufriedenstellend. Die Zuverlässigkeit des Testes ist etwas besser als die der Botanikklausur. Es könnte dies eventuell darauf zurückgeführt werden,

daß die Aufgaben überwiegend bessere Trennschärfeindices aufweisen (van NAERSSSEN 1968). Im übrigen ist der Koeffizient aus denselben Gründen wie bei dem botanischen Leistungstest als unzureichend anzusehen, d.h. beide Leistungstests funktionieren schlecht als Meßinstrument (HUBBARD 1968). Auch hierbei ist aus denselben Gründen wie bei der Botanikklausur die Validität sehr niedrig.

Aus der Interkorrelationsmatrix geht auch hier hervor, daß die wenigen und vorwiegend sehr niedrigen Iteminterkorrelationen für eine gewisse Heterogenität der Klausur sprechen. Dennoch soll im folgenden eine Erklärung der Ergebnisse versucht werden. Die Hauptkomponenten einer Analyse mit 3 extrahierten und rotierten Faktoren repräsentiert im ersten Faktor eine allgemeine zoologische Wissensdimension. Es liegt hier ein "common factor" (ÜBERLA 1968) vor, der mit 8 Items geladen ist, die inhaltlich sowohl entwicklungsgeschichtliche Themen sowie Vererbungslehreprobleme und "reine Zoologie" enthalten. Der zweite Faktor enthält Aufgaben aus dem Bereich Vererbungslehre-Entwicklungsgeschichte, die relativ leicht waren. 3 Aufgaben laden den Faktor negativ, einer positiv. Sie gehören inhaltlich einer Dimension an, messen jedoch etwas Verschiedenes. Worauf der Unterschied beruht, ist nicht erkennbar. Der dritte Faktor ist wieder ein "common factor", dessen inhaltlich den zwei Gruppen zugehörigen Aufgaben den Faktor entgegengesetzt laden. So sind 4 Ladungen von Aufgaben aus der Entwicklungsgeschichte-Vererbungslehrebereich positiv, zwei Aufgaben sogenannter "reiner Zoologie" dagegen laden deutlich negativ. Der Faktor beinhaltet also auch eine allgemeine zoologische Wissensdimension, unterscheidet jedoch durch sein Ladungsmuster zwei verschiedene Bereiche. 4 Items sind bemerkenswerterweise in diesen 3 Dimensionen nicht vertreten. Es fällt auf, daß eine Aufgabe (Nr. 11, s. Anhang) dabei vom testkonstruktorischen Gesichtspunkt nach HUBBARD schlecht ist, ein weiteres Item durch extreme Schwierigkeit und entsprechend schlechte Itemstatistik auffällt, und 2 Aufgaben gehören dem Beantwortungstyp der Weilaufgaben an. Gerade letztere sind in beiden Multiple-Choice-Klausuren deswegen bemerkenswert, weil sie auf dem 5 % Signifikanzniveau keine

Item-Interkorrelationen aufweisen. Außerdem bestehen keine bemerkenswerten Beziehungen zu anderen Variablen dieser Testbatterie, vor allem nicht zu den Items des Intelligenztestes. Dieser Antworttyp soll jedoch besonders dann angewandt werden, wenn schwer erfaßbare Probleme eines Faches angesprochen werden (HUBBARD 1968). Aufgaben dieses Typs sollen nur dann richtig gelöst werden können, wenn die Fähigkeit vorhanden ist, die Frage aus einem größeren Zusammenhang zu begreifen (HUBBARD 1968; SCHULZ, REMSCHMIDT, PRINZ 1970; DAHMER 1970). Die "Beziehungsaufgaben" der hier vorliegenden Klausuren hätten größtenteils besser in anderer Frageform gebracht werden sollen, da sie obigen Kriterien nicht ausgesprochen genügen. Dieser Fragetyp könnte hier eventuell bevorzugt werden sein, nicht weil die Aufgaben grundlegendere, mehr Zusammenhangverständnis fordernde Probleme beinhalteten, sondern weil es an plausibleren Ablenkern fehlte. Gleiches gilt für die Weil-Aufgaben der Botanikklausur.

c) Diskussion der Klausuren: Physik

Die Physik Klausur mit freier Beantwortung jedoch in Form einer "anwendungsbezogenen Kurzantwortform" (DAVE 1968) hat entsprechend ihrer Aufgabeninhalte gute Punktwerte eingebracht. Der MW entspricht 49 % aller erreichbaren Punkte mit jedoch bemerkenswert großer Streuung, als Zeichen einer stärkeren Fähigkeitsstreuung der Bewerber (SCHRADER 1968 b), oder einer besseren Diskriminierung. Die Rohwertverteilung ist relativ symmetrisch, gewiß eine Folge der Aufgabenschwierigkeiten (HUBBARD 1968) der fast nur leichten und mittelschweren Subtests. Die geschätzte Zuverlässigkeit dieses Testes ergibt einen sehr hohen Wert. Man kann ihn jedoch nicht ohne weiteres mit den Reliabilitätskoeffizienten der Multiple-Choice-Klausuren vergleichen, da einerseits zwei völlig verschiedene Testformen vorliegen, andererseits die hier ermittelten Zuverlässigkeiten nach einer anderen Methode berechnet werden. Der hier vorliegende Wert stellt eine echte Korrelation zwischen den Leistungen in zwei Hälften des selben Tests dar, die hier außerdem aus jeweils nur 5 Aufgaben bestehen, so daß der hohe Wert verständlich wird.

Das Interkorrelationsmuster der Klausur zeigt auch hier, daß es sich um einen heterogenen Test handelt. Einige Aufgaben fallen durch zahlreichere Beziehungen zu anderen Aufgabengruppen auf, die sich vielfach inhaltlich erklären lassen. So z.B. der erste Subtest, der in vielen Unteraufgaben hauptsächlich Angaben von physikalischen Definitionen und Einheiten erfordert, was in anderen Aufgaben auch verlangt wird. Ein weiteres Beispiel bietet Subtest 9, der über den "Aufbau der Materie" befragt und mit den Aufgaben "radioaktiver Zerfall", mit "Erhaltungssätze" und anderen mit diesem Problemkomplex in Beziehung stehenden Aufgaben korreliert.

Noch deutlicher kommen diese Wechselbeziehungen in der internen Faktorenanalyse zum Ausdruck. Betrachtet man den Varianzanteil, der bei dieser HKA im Vergleich zu den Multiple-Choice-Klausuren erklärt wird, so fällt auf, daß er bedeutend größer ist. Da bei Tests mit freier Beantwortung die Punktwolken eher als bei Richtig-Falsch-Aufgaben einem Ellipsoid ähneln, ist es einfacher, durch Varimax-Rotation Achsen zu erhalten, die so durch die Punktwolken gehen, daß sie ein Maximum an Varianz erklären. Außerdem ist hier das Verhältnis von Variablenzahl und Faktoren günstiger. Das Ergebnis dieser HKA zeigt, daß die 1. Dimension von den drei Subtests "Arbeit und Leistung", "Kreisbewegung", "Beschleunigte Bewegung" gebildet wird. Man könnte ihn vielleicht "Mechanikfaktor" bezeichnen. Die Varianz des zweiten Faktors wird dagegen aufgeklärt von den drei Subtests "radioaktiver Zerfall", "Erhaltungssätze" und "Aufbau der Materie" und entspricht somit einer Dimension "Atomphysik und Energie". Der dritte Faktor erfaßt, und zwar mit gegensätzlichen Ladungen, zwei als besonders schwierig empfundene Items. Der Subtest "Kausalität" lädt den Faktor negativ und jener, der mit "Viskosität" bezeichnet wird, positiv. Die unter dem Überbegriff "Kausalität" zusammengefaßte Aufgabengruppe zeichnet sich inhaltlich dadurch aus, daß sie im Rahmen dieses Leistungstestes vielleicht als fremdartig empfunden wurde, da nicht unmittelbar physikalisches Wissen erfragt wird. Der Subtest "Viskosität" dagegen wurde als besonders schwierig angesehen, zeichnet sich aber durch gute Trennschärfe aus. Er wurde hauptsächlich nur von guten Probanden gelöst. Im Inter-

korrelationsmuster fällt bei der "Viskositäts"-Aufgabe auf, daß sie keine korrelative Bezüge zu anderen Variablen besitzt. Beide sind jedoch, wenn auch gegensätzlich, durch ihre Dimensionszugehörigkeit miteinander verknüpft. Bezeichnenderweise auf keinem der drei Faktoren vertreten sind die Subtests Nr. 5 und 6 (s. Anhang), in denen es allgemein um "Größen und Einheiten" geht. Sie erfassen keine eigene von den anderen Faktoren deutlich abgrenzbare Dimension, da ihre Einzelaufgaben in vergleichbarer Art auch als Aufgabenkomponenten in den anderen Lösungseinheiten enthalten sind.

Als Beispiel einer typischen Aufgabe dieser Klausur sei der erste Subtest vorgestellt. Er ist zwar umfangreicher als die übrigen, da er aus 11 Items besteht, aber in seiner Aufgabenstellung in kurzen zumeist im Imperativ gebrauchten Worten charakteristisch für diese Klausur. Die Antwortmöglichkeit ist festgelegt, da nur eine bestimmte Formel erfragt wird. Durch ergänzende Erläuterungen konnten Zusatzpunkte erworben werden. Von den bei dieser Aufgabe maximal 14 erreichbaren Punkten wurden im Mittel 5,5 mit einer Streuung von $\sigma = 2,7$ erreicht. Der Schwierigkeitsindex bzw. dessen Äquivalent von 39,2 % zeigt an, daß es sich um eine mittelschwere Aufgabe handelt, mit einem außerordentlich guten Diskriminationsvermögen, da das Trennschärfeindexäquivalent 0,78 beträgt. Der Subtest zeichnet sich aus durch besonders zahlreiche Korrelationen mit anderen Variablen, so z.B. zu den Klausuren Chemie mit $r = 0,35$, Zoologie mit $r = 0,39$ und besonders zu dem I-S-T-AMTHAUER-Item "GE" ($r = 0,31$), das die Abstraktionsfähigkeit prüft. Obwohl die Aufgabe nur Definitionen, also reines Faktenwissen zu erfordern scheint, kann man aus den Korrelationen doch entnehmen, daß sie mehr als nur dieses prüft.

Durch die Besonderheit der Klausur als Test mit freier aber sehr beschränkter Beantwortungsmöglichkeit verringern sich die Vorteile die "Essay-Tests" nachgesagt werden, aber auch deren Nachteile. Als wesentlicher Vorzug dieser Testart gilt, daß sie vom Prüfling ein exaktes Durchdenken eines komplexeren Problems und die Mobilisierung eigenen Wissens erfordert, mit anschließender exakter und stilistisch möglichst guter schriftlicher Darstellung (HUBBARD 1968). Der bedeutendste Nachteil dieses Prüfungstypes, die nachgewiesenermaßen mangelnde

Bewertungsobjektivität (HARTONG 1935; VERNON und MILLICAN 1954) wird hierbei natürlich entscheidend verringert, jedoch nicht ganz ausgeschaltet. Da mehrere Personen die Arbeiten auswerteten, kam es zum Teil von Benotungsunterschieden von einem Punkt pro Aufgabe.

So wie dieser Test hier konzipiert ist, wäre es analog den von MOORE (1954) erarbeiteten Beispielen nicht allzu schwierig, diese Aufgaben in Multiple-Choice-Items umzuformen.

d) Diskussion der Chemieklausur

Die Chemieklausur als zweiter Test mit freier Beantwortung ist die schwierigste unter den 4 naturwissenschaftlichen Prüfungen. Der mittlere Punktwert mit 35 % der erreichbaren Punkte sowie der MW der SI zeigen dieses an. Die überwiegende Zahl der Aufgaben ist äußerst schwer, keine einzige liegt oberhalb des mittleren Schwierigkeitsgrades von 50. Entsprechend zeigt die Rohwerteverteilung ein Maximum in den Klassen niedriger Punktwerte mit deutlich abfallender Tendenz zum Bereich höhere Punkte (Abbildung 5). Sie ist daher auch auf höheren als 1 % - Signifikanzniveau nicht normalverteilt. Für die Zuverlässigkeit dieses Testes gilt das gleiche wie für die oben beschriebene Physikklausur. Auffällig sind die sehr guten Trennschärfeindizes dieser Aufgaben, besonders im Vergleich zu den Multiple-Choice-Klausuren. Ursache dafür könnte einmal die Testart mit freier Beantwortung sein. Die Beziehung zwischen SI und TI (LIENERT 1967) gilt nur für objektive Testverfahren. Da hier keine Antworten vorgegeben sind, kann ein SI von 20 nicht allein durch Zufallstreffer zustande kommen. Aufgaben solcher Schwierigkeit können hier dagegen nur von sehr guten Probanden richtig gelöst werden, so daß gute Trennschärfen bei sehr schweren Aufgaben verständlich werden. Da hier ein ganz umschriebenes Wissensgebiet überprüft wird und die Klausur, wie sich als wichtiges Ergebnis und im Unterschied zu den anderen Klausuren aus dem Interkorrelationsmuster entnehmen läßt, sehr homogen ist, sind die Aufgaben alle etwa gleich schwer, und ein Pbn kann sie entweder alle gut lösen oder nicht. Außerdem wird aus dem Interkorrelationsmuster ersichtlich, daß

die Aufgaben sehr zahlreich und in etwa gleichen Größenordnungen korrelieren, was gewiß eine Folge der auch inhaltlichen Homogenität der Klausur ist. Die interne HKA zeigt, daß der erste Faktor von besonderer Wichtigkeit ist, da er allein 50 % der Varianz erklärt. Er wird von den Aufgaben 2, 3, 5 und 7 geladen. Warum die Lösung dieser Aufgaben im Rahmen der gesamten Klausur von so besonderer Wichtigkeit gewesen zu sein scheint, ist inhaltlich wegen der Homogenität der Aufgaben nicht zu interpretieren.

e) Wechselbeziehungen der Klausuren mit externen Variablen

α) Klausuren und I-S-T-AMTHAUER

Eine Analyse dieser Korrelationen war uns ein besonderes Anliegen, da den objektiven Leistungstests häufig vorgeworfen wird, sie prüften nur Faktenwissen und es brauche lediglich die vorgegebene richtige Lösung gesucht zu werden (MOORE 1956). Dagegen steht die Behauptung, daß diese Testverfahren in ihren Leistungen weit darüber hinaus gehen (HUBBARD 1968; HÖHN 1969; MACINTOSH 1968; PATERSON 1926; DAHMER 1970; SCHULZ, REMSCHMIDT, PRINZ 1970). Im amerikanischen Schrifttum finden sich besonders viele Untersuchungen, welche die Testformen mit freier Beantwortung und die modernen Leistungstests mit Mehrfachwahlaufgaben und ihre jeweilige Beziehung zur Intelligenz analysieren und vergleichen. So kommen fast alle Autoren zu dem Ergebnis, daß die "objektiven" Testverfahren nicht nur objektiver und reliabler sind, sondern auch höhere Ansprüche an die Intelligenz stellen (HUBBARD 1968). Übereinstimmend beweisen z.B. EURICH (1931) und FASTIER (1959), daß Multiple-Choice-Tests signifikant höher mit der Intelligenz korrelieren, als anderen Testverfahren. Besonders Organisationsfähigkeit und Selektionsvermögen (FASTIER 1959) und Verständnis für Regeln und Differenzierungsmöglichkeit (DAHMER 1970) sollen entscheidend mitgeprüft werden.

Wir korrelierten die naturwissenschaftlichen Leistungen der Bewerber mit ihren Intelligenztestergebnissen (Abbildung 31) und fanden, daß alle vier Tests in bestimmter Weise mit den im AMTHAUER gemessenen Aspekten der Intelligenz in Zusammenhang stehen. Im Unterschied zu den in amerikanischen Arbeiten be-

schriebenen Ergebnissen ergibt sich aus der hier vorliegenden Analyse, daß unsere Tests mit freier Beantwortung nicht wesentlich mehr mit der Intelligenz korrelieren, als unsere Multiple-Choice-Tests (Abbildung 19). Die Unterschiedlichkeit der Ergebnisse ist wahrscheinlich in der Verschiedenartigkeit der jeweils analysierten Tests zu suchen. Im einzelnen findet sich (Abbildung 19), daß unser Leistungstest in Botanik besonders die Fähigkeiten der Abstraktionsfähigkeit (GE), des induktiven Denkens mit Zahlen (ZR) und der allgemeinen Vorstellungsfähigkeit (FA) erforderte. Mit dem Intelligenzquotienten (IQ) der Bewerber korreliert dieser Test im Vergleich zu den anderen am höchsten. Auch die Zoologieklausur machte besonders induktives Denken mit Zahlen (ZR) erforderlich, vor allem aber waren die Leistungen dieses Tests von der allgemeinen Merkfähigkeit (ME) abhängig. Die schriftliche Physikprüfung verlangte praktisch-rechnerisches Denken (RA) mehr als andere Intelligenzmodalitäten, prüfte also hauptsächlich die mit dem I-S-T-Item RA ebenfalls gemessene "Fähigkeit des allgemeinen Durchdenkens" (PRIESTER 1964). Darüber hinaus zeigt sich, daß die Abstraktionsfähigkeit der Bewerber in signifikantem Zusammenhang mit den Leistungen in der Physikklausur stehen. Die Ergebnisse der "frei" beantwortbaren Chemieprüfung dagegen sind, außer von der allgemeinen Intelligenzhöhe (IQ) der Bewerber, nur von deren Merkfähigkeit beeinflusst.

3. Beziehungen Klausuren/Abitur

Da die Klausuren bei diesem Studentenkollektiv mit schlechten Abiturzeugnissen über die Zulassung entscheiden sollten, interessierte uns die Frage, ob die Tests nun eine neue allein studienbezogene Leistung messen oder ob Beziehungen zwischen diesen Leistungskomplexen bestehen. Über eine solche zwischen Multiple-Choice-Tests und Schulleistungen (school ratings) liegen Beobachtungen von HUBBARD und CLEMANS (1960) vor. Diese Autoren fanden hohe Korrelationen zwischen Schulleistungen und Ergebnissen objektiver Tests. So eindeutige Beziehungen konnten wir bei unseren Untersuchungen nicht feststellen. Bei r-Werten um 0,3 (Tabelle 18) liegen die Korrelationen nur gerade eben über der Signifikanzgrenze. Die selbe Beziehung ließ sich aber auch

zwischen den MP und den Klausuren mit freier Beantwortung nachweisen. In dieser Hinsicht ließ sich also zwischen beiden Prüfungsarten und deren jeweiligen Beziehungen zu den Schulleistungen bei uns kein Unterschied feststellen.

Die Ergebnisse der einzelnen Subtests der Physikklausur korrelieren überhaupt nicht mit den MP, die Aufgaben der Chemieklausur dagegen sehr stark. Dieser auffallende Unterschied der Korrelationen der beiden von der Konstruktion her gleichartigen Klausuren mit dem MP ist insofern verwunderlich, da zumindest vergleichbare Beziehungen der Physikklausur einerseits und der Chemieklausur andererseits mit den MP erwartet wurden, weil beide Schulfächer gleichermaßen in den MP repräsentiert sind. Die Tatsache, daß kein einziges Physikitem, dagegen 7 Chemieaufgaben mit den MP korrelieren, könnte in dem unterschiedlichen Charakter der Klausuren seine Begründung finden. Während der Physiktest heterogen ist, zeichnet sich die Chemieklausur durch ausgesprochene Homogenität aus.

γ) Beziehungen Klausuren/Vorphysikum

Wenn man bei der Konzeption der Klausuren davon ausgegangen war, daß ein gutes Abschneiden in dieser Prüfung eine gewisse Vorhersage auf den zukünftigen Studienerfolg in demselben Fach zuließ, dann ist der Vergleich der Klausurergebnisse des Vorphysikums von besonderem Interesse. Als einziges erweist sich jedoch nur bei dem Fach Physik die Beziehung zwischen Klausur und Vorphysikumsprüfung als signifikant. Dieser Klausur kommt demnach eine gewisse prognostische Bedeutung zu. In den übrigen 3 Fächern ist nach unseren Untersuchungen eine solche Prognose nicht sicher möglich.

2. Leistungskomplex Abitur

a) Beziehungen der Noten zum Vorphysikum

Wie bereits oben erwähnt (Kap.IV, 2), sind sich viele Experten einig, daß Schulnoten ein ungeeignetes Selektionskriterium darstellen. Diese Erfahrungstatsache läßt sich auch aus unseren Ergebnissen schließen, insbesondere durch die niedrigen Korrelationen zwischen Abiturnoten insgesamt und den Studienleistungen. Im einzelnen zeigen unsere Untersuchungen, daß die Note im Fach Deutsch keine Beziehung zu den Studienleistungen des Vorphysikums aufweist. Zu demselben Ergebnis kommt auch SCHÜTZ, E. (1968), der darüber hinaus gleiches für das Physikum feststellte. Während bei unserem Kollektiv kein Zusammenhang zwischen Mathematik und dem Vorphysikum besteht, findet SCHÜTZ gerade hier interessanterweise mit $r = 0,29$ den höchsten Koeffizienten zwischen den Leistungen in der Schule und im Studium. Geht man jedoch von den Klausuren als vergleichbaren Studienleistungen aus, so findet sich auch bei unseren Untersuchungen - wie bereits oben gesagt - eine ähnlich enge Beziehung (Abb. 31). Die Abiturnote in Physik unserer Pbn hat keinen Einfluß auf den späteren Studienerfolg - repräsentiert durch das Vorphysikum. SCHÜTZ beschreibt hier einen Koeffizienten um 0,2 dieser Abiturnote mit den Vorphysikumsprüfungen in Physik und Chemie. Dagegen zeigt sich bei uns, daß die Abiturnote in Chemie für die Leistung desselben Faches im Vorphysikum Prognosen erlaubt. Beide Noten korrelieren relativ hoch miteinander (Abb. 31) und sind in einer Hauptkomponentenanalyse mit ihren Ladungen im 4. Faktor enthalten (Tab. 33).

Es fällt auf, daß die Abiturzensur in Gemeinschaftskunde relativ hoch mit den Prüfungen in Zoologie des Vorphysikums korreliert. Wie sehr diese Zensuren vergleichbare Leistungen vorauszusetzen scheinen, wird auch in einer HKA (Abb. 36) verdeutlicht, bei der diese Noten zusammen mit Botanik gemeinsam die 3. Leistungsdimension bilden. SCHÜTZ findet keine so deutliche Beziehung zwischen Gemeinschaftskunde und den Vorphysikums- und Physikumsleistungen. Man muß allerdings berücksichtigen, daß er im Unterschied zu uns die Zensur in

Gemeinschaftskunde mit den Prüfungsendzensuren in Vorphysikum und Physikum korrelierte, die Summenzensuren darstellen.

Eine Korrelation der Kernfächerzensur des Abiturs (MP) mit der Vorphysikumsgesamtzensur (Abb. 30) ergab bei uns, daß keine Beziehung zwischen Abitur- und Vorphysikumsgesamtnote vorhanden ist. Die Selektion nach den MP leistete also keine Trennung der Pbn mit potentiell besseren Studienleistungen von voraussichtlich schlechteren Studenten.

b) Beziehungen der Abiturnoten zueinander und Vergleiche mit der Literatur

Wie BOBERTAG, O. (1915) und PFAFF, A. (1961) darstellen, bestehen verschiedene interessante Beziehungen zwischen den Leistungen der Schulfächer. So finden z.B. beide übereinstimmend hohe Korrelationskoeffizienten des Faches Deutsch mit Erdkunde und Geschichte. Auch bei uns findet sich (Abb. 25) gleiches zwischen diesem und dem Fach Gemeinschaftskunde, in welchem Erdkunde und Geschichte enthalten sind. Bei uns läßt sich darüber hinaus diese gefundene deutliche Beziehung in verschiedenen HKA noch eindrucksvoller zum Ausdruck bringen. Wie aus Abb. 34 hervorgeht, bilden beide Fächer eine eigenständige, von allen anderen Variabeln deutlich isolierte Leistungsdimension. Besondere Beziehungen zeigen sich bei uns auch zwischen den Noten in Deutsch und Biologie. Auch BOBERTAG findet einen hohen Koeffizienten von 0,63 in diesen Fächern. Das Ergebnis von PFAFF liegt mit einem r-Wert von 0,29 nur geringfügig unter dem von uns erhaltenen Koeffizienten. Während beide Autoren, außerdem auch WEINGARDT, E. (1964) besonders starke Beziehungen zwischen den Fächern Deutsch und Englisch finden, läßt sich ähnliches bei uns nicht feststellen. Die Abiturnoten in Mathematik zeigen bei uns besondere Korrelationen mit naturwissenschaftlichen Fächern wie Physik und Chemie. Auch BOBERTAG und PFAFF finden die höchsten Korrelationswerte der Mathematikzensur mit diesen beiden Schulfächern. Wie ausgeprägt die Wechselbeziehungen gerade dieser drei Fächer sind, kommt bei uns darüber hinaus in einer HKA (Abb. 33) zum Ausdruck, wo diese Abiturfächer mit

ihren Ladungen in einem naturwissenschaftlichen Leistungsfaktor enthalten sind. Bemerkenswerterweise zeigen die mit den oben beschriebenen Schulfächern ebenfalls in Beziehung gesetzten Noten in Sport und Englisch keinerlei signifikante Korrelation zu denselbigen.

Interessanterweise findet man bei einem Vergleich der mittleren Abiturleistungen unseres Kollektivs mit den von WEINGARDT berechneten Mittelwerten und Streuungen der Zensuren von Abiturienten, daß unsere Pbn in den Fächern Deutsch und Mathematik etwas schlechtere Zensuren aufweisen, als der wohl repräsentativere Querschnitt von WEINGARDT. In Chemie, Physik und Englisch sind die Leistungen beider Gruppen gleich, in Biologie zeigen unsere Bewerber etwas bessere Leistungen als der allgemeine Abiturientendurchschnitt.

c) Beziehungen zwischen Schulleistungen und Intelligenz

Inwieweit Schulleistungen mit der allgemeinen Intelligenz, wie sie z.B. im I-S-T-AMTHAUER gemessen wird, in Beziehung stehen, ergibt sich aus Abb. 31. Es zeigt sich, daß Deutsch, besonders mit der allgemeinen Fähigkeit der "Urteilsbildung" (SE), (AMTHAUER, R. 1955) korreliert, Mathematik mit dem "induktiven Denken in Zahlen" (ZR). Letztere Beziehung kommt besonders deutlich auch darin zum Ausdruck, daß in der HKA (Abb. 33) beide den ersten Faktor laden. Interessant ist, daß dieses Schulfach nicht mit RA, also Rechenaufgaben korreliert, wie man vielleicht erwarten könnte. Auf diese empirisch gefundene Tatsache weist auch GUILFORD, J.P. (1965) ganz besonders hin. Die Schulphysik dagegen zeigt keine wesentlichen Bezüge zur Intelligenz. Die Chemie korreliert mit dem "Erfassen von Bedeutungsgehalten" und der "räumlichen Vorstellungsfähigkeit". Besonders letztere Fähigkeit scheint für Chemie von besonderer Bedeutung, da beide in einer HKA (Abb. 34) einen eigenen Faktor bilden. Das Fach Gemeinschaftskunde erfordert anscheinend besonders allgemeine Intelligenzmodalitäten, so z.B. "Abstraktionsfähigkeit" (Abb. 31) und "räumliche Vorstellungsfähigkeit". Nach GUILFORD (1965) sollen besonders die verbalen Intelligenztestitems mit Schulleistungen zusammenhängen.

3. Leistungskomplex: Vorphysikum

Außer den oben genannten Beziehungen zwischen dem Vorphysikum und den Abiturleistungen und der bereits erwähnten einzigen Korrelation des Physiktestes (aus dem gesamten Klausurkomplex) mit der mündlichen Vorphysikumsprüfung ist es interessant zu wissen, ob in diesen Prüfungen besondere Intelligenzleistungen erforderlich sind. Während beide Leistungskomplexe nur recht schwach korrelieren und auch in der HKA (Abb. 33) deutlich wird, daß naturwissenschaftliche Leistungsdimensionen praktisch gar nichts mit den Intelligenzdimensionen gemein haben, zeigt sich in der Korrelationsanalyse, daß in den einzelnen Vorphysikumsprüfungen zumindest einige Intelligenzvariable mitgeprüft werden.

Besonders im Fach Zoologie ergeben sich Beziehungen zur allgemeinen Merkfähigkeit, zu praktisch-rechnerischem Denken und der Abstraktionsfähigkeit. Letzteres wurde auch in der Physikprüfung gefordert, wogegen rechnerische Fähigkeiten scheinbaren Einfluß auf die Ergebnisse der Chemieprüfung hatten.

Es bleibt jedoch zu diskutieren, ob die im I-S-T-AMTHAUER gemessenen Fähigkeiten, die ja unter normalen emotionalen Bedingungen gemessen werden sollen, um als echte Intelligenzleistung angesehen werden zu können (GROFFMANN, K.J. 1964, 1963), mit denen einer Prüfungssituation so ohne weiteres gleichgesetzt werden können.

In den Hauptkomponentenanalysen (Abb. 34) fällt auf, daß die mündliche Prüfungsreihe des Vorphysikums hauptsächlich in zwei Gruppen zerfällt, welcher einerseits die Botanik- und Zoologieleistung zugehörig ist, andererseits die Leistungen der Physik- und Chemieprüfung kombiniert auftreten. Erstere sind gewöhnlich in einer eigenständigen Dimension vertreten, letztere dagegen bilden zusammen mit der Physik Klausur einen eigenen Faktor. Möglicherweise bestehen unterschiedliche Anforderungen zwischen der Biologiegruppe einerseits und der Physik-Chemiegruppe andererseits.

4. Leistungskomplex: I-S-T-AMTHAUER

a) Beziehungen zwischen Intelligenz und den übrigen Leistungskomplexen

Die Bedeutung des I-S-T im Rahmen dieser Testanalyse wurde bereits diskutiert. Abgesehen von den detaillierten Beziehungen der einzelnen Klausuren, der Abitur- und Vorphysikumsnoten mit einzelnen Intelligenzuntereinheiten, zeigen die größeren Zusammenhänge in den Leistungen der 4 Gesamtkomplexe im wesentlichen, daß die Testleistung im I-S-T keine starken Beziehungen zu den Abiturnoten wie auch zur Studienleistung in Form der mündlichen Vorphysikumsprüfung aufweist. Die bei unserem Kollektiv nur schwach vorhandene Korrelation zwischen Abiturnoten und dem I-S-T könnte die Auffassung von JUST (1939) stützen, der aus seinen Ergebnissen schloß, daß "Ärzte einem in der Schulleistung nicht unmittelbar meßbaren und zum Ausdruck kommenden Begabungstypus entsprechen". Unsere Ergebnisse bestätigen somit auch die von HELEN, J. (1967) analysierten Beziehungen zwischen Studienleistungen und IQ, die er als unbedeutend fand. HELEN deutet die mangelnde Beziehung zwischen Intelligenz und Studienleistungen gerade der ersten Studienjahre mit der Annahme, daß zu Studienbeginn Motivationsfaktoren, wie Hoffnung auf Erfolg oder Furcht vor Mißerfolg, für die Leistungsvarianz bedeutender sind als die Intelligenzunterschiede der Pbn. Da sich für unser Studentenkollektiv in seiner besonderen Situation ähnliche Motivationsfaktoren ausgewirkt haben könnten, wären die gefundenen Beziehungen u.U. dadurch erklärbar.

b) Diskussion der Intelligenzprofile

Über die Analyse obiger Beziehungen hinausgehend, interessierte uns ein Weiteres. Unter der Voraussetzung, daß die Hypothese AMTHAUERS richtig ist, daß nämlich eine bestimmte, berufstypische Intelligenzstruktur unabhängig von der Berufsausübung, also auch schon vorher, determinativ im Menschen angelegt ist, interessierten wir uns dafür, ob unsere Pbn ein solches Berufsprofil aufweisen. Im positiven Fall könnten sie,

nach AMTHAUER zumindest, voraussichtlich gute Mediziner und nicht nur den Leistungen nach gute Medizinstudenten werden. Ihr Intelligenzprofil wurde ermittelt und mit dem "Ärzteprofil" AMTHAUER's verglichen (Abb. 26). Das Ergebnis zeigt die Ähnlichkeit beider Intelligenzmuster. Das Profil der Studenten zeigt nur Abweichungen in den Subtests WA, RA und ZR, die bedingt sein dürften durch den noch seitens der Schule vorhandenen Übungseffekt unserer Pbn.

Unsere Leistungstests betreffend interessierte außerdem die Beantwortung der Frage, ob sich die Intelligenzprofile der Pbn mit überdurchschnittlichen Leistungen in den Klausuren, das sind Bewerber, die später zugelassen wurden, deutlich unterscheiden von denen mit unterdurchschnittlichen naturwissenschaftlichen Klausurergebnissen. Das Kollektiv wurde in entsprechende Gruppen eingeteilt, und ihre zugehörigen Profilverläufe verglichen (Abb. 27). Würden nämlich die Verläufe signifikant unterschiedlich sein, könnte man den Klausuren Diskriminationsfähigkeit zuschreiben. Sie hätten dann die Fähigkeit, einzelne gute von weniger intelligenten Studenten zu trennen, bzw. in Anlehnung an AMTHAUER prospektiv gute Mediziner aus einer Pbn-Gruppe zu selektieren. Der Profilvergleich von über- und unterdurchschnittlichen Klausurteilnehmergruppen zeigt jedoch keine charakteristischen Unterschiede (Abb. 27). Die Pbn mit guten Klausurergebnissen lagen im ganzen gesehen mit ihrem Profilniveau etwas höher als das Ärzteprofil nach AMTHAUER, während die schlechteren Pbn etwas darunter lagen.

Eine weitere Untersuchung befaßt sich mit der Frage, ob die Leistungen nach bereits zweiseimestrigem vorklinischen Studium der Medizin anhand der Vorphysikumsnoten eine ähnliche Diskrimination in intelligente bzw. nach AMTHAUER geeignete Pbn, von weniger intelligenten erlaubt. Es wurden auch die Intelligenzprofile der Studenten, die nach erreichter Aufnahme zum Medizinstudium nach über- und unterdurchschnittlichen Vorphysikumszensuren getrennt worden waren, verglichen (Abb. 28). Das Ergebnis ist dem obigen vergleichbar. Die Profilstrukturen

unterscheiden sich nicht wesentlich. Es zeigt sich jedoch, daß die Pbn der unterdurchschnittlichen Vorphysikumsgruppe deutlich schwächere Leistungen in den Items GE und ZR aufweisen als die besseren Vorphysikumskandidaten. Da dieser Unterschied auch statistisch signifikant ist, läßt sich feststellen, daß die guten Vorphysikumsprobanden über bessere Fähigkeiten abstrakten und theoretisch-rechnerischen Denkens verfügen. Eine ähnliche Bedeutung guten logischen Denkvermögens und theoretisch-rechnerischen Denkens für die Leistungen in den Klausuren wurde oben bereits angedeutet.

c) Die Hauptkomponentenanalyse des I-S-T brachte als Ergebnis 4 Faktoren, die etwa folgendermaßen bezeichnet werden können:

1. Faktor: Dimension des mathematischen Denkens mit Kombinations- und Abstraktionsfähigkeit,

2. Faktor: Räumliche Vorstellungsfähigkeit,

3. Faktor: Dimension der allgemeinen Vorstellungs- und Merkfähigkeit,

4. Faktor: Sprachlicher Faktor.

Es lassen sich Gruppen von Intelligenzmodalitäten isolieren, die in funktionalem Zusammenhang stehen, sich auf gemeinsame Bedingungen zurückführen lassen (JÄGER, 1967). Die Interpretation der erhaltenen Intelligenzfaktoren kann nur mit Vorsicht erfolgen, und ist etwas unsicher (GUILFORD, 1965).

Eine Beschreibung der Faktoren ergibt bei uns folgendes: Im ersten Faktor finden wir die Fähigkeiten "semantischen Beziehungserkennens", des "allgemeinen Durchdenkens" und des "induktiven Denkens" (AMTHAUER, 1955). Er entspricht also im wesentlichen der Fähigkeit der Pbn in numerischen und abstrakten Beziehungen zu denken (logisches Denken). Er scheint für unser Kollektiv und im Zusammenhang mit dieser speziellen Testbatterie von besonderer Bedeutung. Die 2. Dimension entspricht dagegen ausschließlich bildlicher Denkinhalte, dem visuellen Gedächtnis oder auch - nach GUILFORD - einem Faktor des "Ordnen bei konvergentem Denken". Auch im 3. Faktor kommen im wesentlichen bildungsunabhängige Intelligenzleistungen zum Ausdruck, wie das "Erkennen figuraler Beziehungen" (MEILI, 1965) und allgemeine Gedächtnisleistungen, im besonderen das Kurzzeitgedächtnis. Man erhält hierbei auch Hinweise

auf die Konzentrationsfähigkeit der Pbn (AMTHAUER, 1955). Der 4. Faktor, der sprachliche Fähigkeiten voraussetzt, ist bildungsabhängig. Hier kommen hauptsächlich semantische Denkinhalte zum Ausdruck, auch die Fähigkeiten des Erkennens und sinngemäßen Reproduzierens von sinnvollen Zusammenhängen. In den hier erhaltenen Faktoren können natürlich nur die Intelligenzleistungen als Dimensionen isoliert werden, die im I-S-T gemessen werden. Sie können daher und auch wegen der besonderen psychischen Situation, aus der die Ergebnisse gewonnen wurden, nur für unser Kollektiv gewertet werden. Dennoch dürfte es interessant sein, unsere Ergebnisse z.B. mit denen von FISCHER (1958) zu vergleichen, der ebenfalls eine FA des I-S-T berechnete, oder weiterhin unsere Intelligenzdimensionen neben denen von JÄGER zu sehen oder auch neben den von ihm zitierten und anderweitig beschriebenen Intelligenzfaktoren. Es wird hier darauf verzichtet, weil es zu weit von der Thematik dieser Arbeit wegführen würde.

5. Faktorenstruktur des Variablenkollektivs

Die Faktorenanalysen wurden bei uns im wesentlichen als Ordnungsinstrument der zahlreichen Leistungsvariablen angewandt, welche sie nach ihrer funktionalen Ähnlichkeit klassifizieren (JÄGER, 1967). Sie führen die Variablen auf ihre gemeinsamen, oftmals verborgenen Beziehungen zurück und offenbaren diese. Es zeigt sich als wesentliches Übersichtsergebnis, daß die naturwissenschaftlichen Leistungen - in Klausuren und Vorphytikum - von besonderer Wichtigkeit waren, und Faktoren, die diesen Leistungen entsprechen, vor denen des Abiturs und der Intelligenz rangierten. Die 4 großen Leistungsgruppen verteilen sich üblicherweise im wesentlichen auf eigene Dimensionen, die den jeweiligen Prüfungskomplexen zugrundeliegenden Anforderungen und Ergebnissen entsprechen. Kritiker der FA würden dieses Ergebnis als erwartungsgemäß bezeichnen und das Argument, daß die Auswahl des Materials das Ergebnis der FA determiniere, gegen ihre Anwendung verwenden. Dagegen läßt sich einwenden, daß bei jeder wissenschaftlichen Untersuchung die Auswahl des Materials und die Versuchsanordnung das Ergebnis in bestimmter Weise beeinflusst (JÄGER, 1967). So finden

sich bei uns die Leistungsgruppen zwar auf bestimmten Faktoren wieder, doch gibt uns die FA Aufschluß darüber, welche Leistungen innerhalb des großen Variablenkollektivs vorrangig waren und welche unbekannten Beziehungen zwischen den Komplexen bestehen.

Im folgenden soll abschließend die Hauptkomponentenanalyse beschrieben werden, in der die 4 Variablen des Klausurkomplexes, die 5 Variablen des Abiturkomplexes, 4 Vorphysikumswerte und 9 I-S-T-AMTHAUER-Werte untersucht werden (Abb. 33).

Das Ergebnis zeigt an erster Stelle einen naturwissenschaftlichen Leistungsfaktor, der mit den 3 schwierigeren Aufnahme-klausuren geladen ist und außerdem mit der Mathematiknote des Abiturs und dem Zahlenreihentest des I-S-T. Die negative Ladung der Mathematikzensur kommt dadurch zustande, daß eine größere Zahl der Note einer schlechten, eine kleinere Zahl aber einer guten Leistung entspricht. Das Vorzeichen ist demnach umgekehrt zu interpretieren. Mathematische Fähigkeiten stehen also mit den Leistungen in 3 Klausuren in besonderem Zusammenhang.

Unser gesamtes Kollektiv zeigt jedoch gerade auf diesem Gebiet unterdurchschnittliche Leistungen, wie die MW der Mathematikzensur und des Items "ZR" des I-S-T zeigen. Interessanterweise läßt sich am Intelligenzprofil der Pbn feststellen, daß die Bewerber mit guten naturwissenschaftlichen Leistungen auch in den Tests "RA", "ZR" des I-S-T deutlich bessere Ergebnisse haben, als der Rest der Bewerber und auch bessere als die von AMTHAUER untersuchten Ärzte. Die durch die FA zum Ausdruck gebrachte Beziehung zwischen mathematischen Denkfähigkeiten und den Leistungen in 3 naturwissenschaftlichen Klausuren könnte also derart sein, daß die Pbn mit besonders gutem logischen Denkvermögen auch die besseren Klausurergebnisse erreichten. Die 2. Dimension dieser Analyse wird ebenfalls durch naturwissenschaftliche Leistungen repräsentiert, und zwar die der mündlichen Vorphysikumsprüfungen in Zoologie und Botanik. Er könnte somit als Biologie faktor bezeichnet werden. Die inhaltlich enge Beziehung beider Fächer wird hier statistisch bestätigt. An dritter und vierter Stelle rangieren interessanterweise zunächst reine Intelligenzfaktoren aus I-S-T-Items. Sie entsprechen ungefähr dem vierten

und ersten Faktor der internen HKA des I-S-T. Es sind dies der Faktor des Beziehungserkennens und der Dimension, die man etwa als die des "logischen Denkens" bezeichnen könnte. Erst dann folgt an fünfter Stelle wieder ein Faktor naturwissenschaftlicher Leistungen, gebildet von den Vorphysikumsprüfungen in Chemie und Physik. Die Vorzeichen der Ladungen von Vorphysikumsnoten sind wiederum umgekehrt zu interpretieren, da es sich um Zensuren handelt. Es zeigt sich, daß die Leistungen in Physik zu denen in Chemie sich wahrscheinlich umgekehrt proportional verhalten, da sie entgegengesetzte Ladungen haben. Gemeinsam setzen sich deutlich ab vom Biologiefaktor derselben Prüfung und dokumentieren, daß sie eine davon unabhängige Vorphysikumsdimension sind. Der Faktor ist aber außerdem mit der Klausur in Physik geladen, woraus zu entnehmen ist, daß Physik Klausur- und -Vorphysikumsprüfung in bestimmter Weise zusammenhängen. Es könnte vielleicht eine Bestätigung für diese Klausur sein, da man ihr dadurch eine prognostische Aussagefähigkeit für die nächstfolgende Studienleistung in demselben Fach zusprechen kann. Sechstens und siebentens finden wir reine Abiturfaktoren. Die Beziehungen Deutsch-Gemeinschaftskunde wurden bereits diskutiert. Die Schulphysik stellt interessanterweise eine völlig unabhängige Leistungsdimension dar. Der 8. Faktor dokumentiert, daß die mündliche Prüfung in Chemie ganz besonders von der Merkfähigkeit der Pbn beeinflusst scheint, wogegen die Schulchemie mehr von der allgemeinen Fähigkeit des räumlichen Vorstellens abhängt, wie der 9. Faktor offenbart. Man könnte diese Unterschiede evtl. dahingehend deuten, daß die Zensur in der Vorphysikumsprüfung Chemie mehr durch die augenblicklich aktualisierbare Wissensfülle, die ja abhängig ist von der Merkfähigkeit, beeinflusst wird, als dieses bei der Abiturnote in Chemie, die eher ein langjähriges Leistungsergebnis darstellt, der Fall ist. Letztere dagegen zeigt, daß in diesem Fach ganz besonders das Denken in räumlichen Dimensionen geschult wird.

V. Zusammenfassung

Eine Testbatterie aus Abiturnoten, Intelligenz-Struktur-Test, 4 Klausuren und Vorphysikumsnoten wurde an Hand von Korrelations- und Faktorenanalysen untersucht. Es sollten so die im SS 1968 in Marburg verwandten und zum Wechsel vom naturwissenschaftlichen zum Medizinstudium berechtigenden Auswahlkriterien auf ihre Eignung als Selektionsmethode geprüft werden. Es ergaben sich dabei 3 verschiedene Fragenkomplexe:

I. Inwieweit genügen die Klausuren den theoretischen Forderungen als Testinstrument?

a) Die Botanikklausur erweist sich als ungeeignet, da ihr Schwierigkeitsindex zu niedrig, d.h. die Klausur zu schwer ist (zu stark biochemielastig), da auf Grund eines zu niedrigen Validitätskoeffizienten die Klausur nicht das mißt, was sie messen sollte,

da der zu niedrige Reliabilitätskoeffizient auf eine zu geringe statistische Zuverlässigkeit des Testes hinweist, da die Anzahl von 20 Aufgaben für einen Multiple-Choice-Test zu gering ist und schließlich da die Mehrzahl der Aufgaben ungünstige Schwierigkeits- und Trennschärfeindices aufweisen.

b) Die Zoologieklausur ist von mittlerer Schwierigkeit. Die Schwierigkeits- und Trennschärfeindices der Aufgaben zeigen insgesamt bessere Itemkonstruktionen an, als im Botaniktest. Bezüglich Validität und Reliabilität gilt das gleiche wie oben. Aus den Iteminterkorrelationsmatritzen und Hauptkomponentenanalysen dieser sowie der Botanikklausur ergibt sich eine gute inhaltliche Heterogenität.

c) Die als Test mit freier Beantwortung durchgeführte Physikklausur ist leicht bis mittelschwer und enthält in der Mehrzahl Aufgaben mit guten Schwierigkeits- und Trennschärfeindices. Als weiteres Positivum besteht auch hier eine weitgehend heterogene Struktur, nachteilig wirkte sich jedoch die Bewertung durch verschiedene Prüfer aus.

d) Die Chemieklausur hat bei weitem den niedrigsten Schwierigkeitsindex, was gleichzeitig die beste Trennschärfe zur Folge hat. Inhaltlich ist sie sehr homogen.

II. Welche Beziehungen bestehen zwischen den Klausurergebnissen und anderen untersuchten Variablen?

- a) Zwischen den Klausurleistungen und den I-S-T-AMTHAUER-Ergebnissen sowie
- b) zwischen jenen und den Abiturnoten bestehen nur kaum signifikante Korrelationen
- c) ebenso finden sich zwischen den Klausurnoten und den entsprechenden Zensuren im Vorphysikum keine sicheren Beziehungen, außer zwischen der Physikklausur und der Vorphysikumszensur in Physik.

III. Welche Beziehungen bestehen zwischen den Leistungskomplexen oder den einzelnen Variablen untereinander und welche Faktorenstruktur ergibt sich für das gesamte Variablenkollektiv?

- a) Es findet sich im wesentlichen nur zwischen der Chemiezensur des Abiturs und des Vorphysikums eine Korrelation von vertretbarer Signifikanz.
- b) Die Hauptkomponentenanalyse des Vorphysikums ergibt einen Biologiefaktor (Zoologie-Botanik) und einen zweiten davon unabhängigen Physik-Chemie-Faktor.
- c) Die HKA des I-S-T-AMTHAUER weist darauf hin, daß die Fähigkeiten des logischen Denkens und der räumlichen Vorstellung von besonderer Bedeutung sind.
- d) Aus der HKA des gesamten Variablenkollektivs ergibt sich eine deutliche Trennung der verschiedenen Leistungskomplexe, wobei die naturwissenschaftlichen Leistungen besonders der Klausuren vor denen des Vorphysikums, der Intelligenz und dann des Abiturs rangieren.

IV. Außerdem zeigt ein Vergleich der Intelligenzprofile unserer Pbn mit dem von AMTHAUER für Mediziner ermittelten einen ähnlichen Verlauf.

LITERATURVERZEICHNIS

- AMTHAUER, R.: a) Intelligenz und Beruf.
Z.f.exper.u.angew.Psychologie, Bd. 1, 1953
b) Intelligenz-Struktur-Test, I-S-T Handanweisung
für die Ausführung und Auswertung.
Verl.f.Psych.Dr.H.C.Hogrefe, Göttingen, 2.Aufl.1955.
- AUMILLER, J.: MMW, Heft 36, 1967, (S. 1874/75).
- BERNART, E.: Schulleistung und Begabung.
E.Reinhardt Verlag, München-Basel 1962.
- BOBERTAG, O.: Korrelationsstatistische Untersuchungen über die
Unterrichtsleistungen der Schüler einer höheren
Lehranstalt.
Z.f.angew.Psychologie 1915, H. 10.
- BULL, G.M.: Examinations.
J.m.Educ.34, 1959.
- CATTELL, R.B.: The Scree-Test for the Number of Factors.
Multivar.behav.Res.1, (245 - 276) 1966.
- COWLESS, J.T., and HUBBARD, J.P.: A comparative Study of
Essay and objective Examinations for Medical
Students.
J.m.Educ. 27, 1952.
- DAHMER, J.: Schriftl.Erfolgskontrollen - Anleitung zur Er-
stellung von Multiple-Choice-Fragen.
Didacta Medica, Bd.1, 1970.
- DAVE, R.: Eine Taxonomie pädagogischer Ziele und ihre Beziehung
zur Leistungsmessung in: Möglichkeiten und Grenzen
der Testanwendung in der Schule.
Pädagogisches Zentrum, Veröffentlichungen Reihe C:
Berichte, Bd. 15, Verl.J.Beltz, Weinheim-Berlin-
Basel 1968 (Herausg.von Ingenkamp, K.-H. u.Marsolek,T.).
- DUNN, S.S.: Verschiedene Perspektiven zu den Zielen der Test-
anwendung.
(Angaben siehe DAVE).

- EURICH, A.C.: Four Types of Examinations Compared and Evaluated.
J.Educ.Psych. 22, 1931.
- FASTIER, F.N.: Consistency of Performance at Examinations.
J.m.Educ. 34, 1959.
- FURCK, C.L.: Das Pädagogische Problem der Leistung in der Schule.
Weinheim 1961.
- GEORG, D.: Schriftliche Tests als Rückinformation im Medizin-
unterricht.
Didacta Medica, Bd. 1, 1970.
- GROFFMANN, K.J.: Entwicklung der Intelligenzmessung.
Diagn., Bd. 6, 1964.
- GUILFORD, J.P.: Fundamental Statistics in Psychology and
Education.
Mc Graw-Hill, New York 1956.
- GUILFORD, J.P.: Persönlichkeit.
Verl.Beltz, Weinheim, Bergstr. 1965.
- HEHLMANN, W.: Wörterbuch d. Psychologie.
Fischer Verlag 1965.
- HENRYSSON, S.: Methoden der Konstruktion und Analyse von Test-
aufgaben.
Pädagogisches Zentrum, Veröffentlichungen Reihe C:
Berichte, Bd. 15, Verl.J.Beltz, Weinheim-Berlin-
Basel 1968 (Herausg.von Ingenkamp, K.-H.u.Marsolek, T.)
- HÖHN, E.: Objektivierung und Kontrolle der Prüfungen.
In: Prüfungen als Hochschuldidaktisches Problem.
Ergebnisse und Materialien eines Expertenseminars
in Hamburg-Rissen. Jan./Febr.1969 abgedruckt in:
Blickpunkt Hochschuldidaktik H. 1 (1969).
- HOFSTÄTTER, P.: Lexikon der Psychologie A - Z.
Fischer Bücherei 1966.
- HUBBARD, J.P., and CLEMANS, W.V.: A Comparative Evaluation of
Medical Schools.
J.m.Educ., 35, 1960.

HEISS, R.: Zum Begriff der Intelligenz.

Diagnostika, Bd. 6, S. 3 ff, 1964.

Informationsblatt über die Richtlinien der Med.Fakultäten der
BRD. Herausgeg.von der Zentralen Registrierstelle
für die Zulassung zum Studium der Medizin und Zahn-
medizin bei der Rheinischen Friedrich-Wilhelm-
Universität Bonn, 1966.

JÄGER, A.O.: Dimensionen der Intelligenz.

Verl.f.Psych.Dr.C.J.Hogreve, Göttingen 1967.

JUST, G.: Erbpsychologie der Schulbegabung.

Hdb.d.Erbbiologie d.Menschen, Bd. 5, Verl.Jul.
Springer, Berlin 1939.

KAPUSTE, H., u.H.NOACK: Leistungstests im Medizinstudium -
Möglichkeiten und Grenzen.

In: Probleme des Medizinunterrichts.

Herausgeg.vonTh.vonUEXKÜLL, Urban u.Schwarzenberg,
Berlin-München-Basel 1968.

KENDALL, M.G., u.A.STUART: The Advanced Theory of Statistics,
Bd. 2, Griffin London 1962, Zweite Auflage.

KLEINERT, H.: Zensur, Zeugnis, Schulzeugnis.

Lexikon der Pädagogik, Bern 1951.

KÜGELGEN, von A.: Auslese beim Medizinstudium.

DÄ, H. 2, 1966, H. 3, 1966, H. 4, 1966.

KAPUSTE, H.: Probleme der Zulassung zum Medizinstudium in der
Bundesrepublik.

Sonderdruck d.Deutschen Ärzteblattes - Ärztl.Mit-
teilungen 1966.

LEE, J.M., and SYMONDS, P.M.: New-Type of Objective Tests:

A Summary of Recent Investigation.

J.Educ.Psych.24, 1933; 25, 1934.

LIENERT, G.A.: Faktorenstruktur der Intelligenz als Funktion
des Neurotizismus.

Z.f.exp.u.angew.Psych., Bd. 10, 1963.

LIENERT, G.A.: Testaufbau und Testanalyse.

J.Beltz, Weinheim-Berlin 1967.

LIENERT, G.A.: Verteilungsfreie Methoden in der Biostatistik.

Verlag Anton Hain, Meisenheim an der Glahn, 1962.

LINDQUIST, E.: Educational Measurement.

American Council on Educ., Washington, D.C. 1957.

MEILI, R.: Lehrbuch der psychol.Diagnostik.

Verlag Huber, Bern-Stuttgart 1965.

MERZ, .: Vorlesungsskriptum.

Different.Psych.d.Univ.Marburg/L.

MITTENECKER, E.: Planung und statistische Auswertung von Experimenten.

Verlag Franz Deutzcke, Wien, 1966.

MOORE, R.A.: Methods of Examining Students in Medicine.

J.m.Educ., 29, 1954.

NAERSSSEN, van R.F.: Ein "Signal/noise ratio"-Index zur Aufgaben-
auslese bei informellen Tests.

Pädagogisches Zentrum, Veröffentlichungen Reihe C:

Berichte, Bd. 15, Verl.J.Beltz, Weinheim-Berlin-

Basel 1968 (Herausg.von Ingenkamp, K.-H.u.Marsolek, T.).

PAWLIK, K.: Dimensionen des Verhaltens.

Verlag Hans Huber, Bern-Stuttgart 1968.

PFAFF, A.: Die Interkorrelationen von Schulnoten an der Ober-
stufe neusprachl.Gymnasien.

Psych.Institut d.Universität Marburg/L. 1961.

PRIESTER, H.-J.: Der Intelligenz-Struktur-Test von AMTHAUER.

Handbuch d.psych.Diagn., Bd. 6, S.240 ff, 1964.

SACHS, L.: Stat.Auswertungsmethoden.

Springer Verlag, Berlin-Heidelberg-New York 1970,

2. Aufl.

SCHRADER, W.B.: Validierungsuntersuchungen und Normen als Hilfs-
mittel bei der Interpretation von Testwerten.

Pädagogisches Zentrum, Veröffentlichungen Reihe C:

Berichte, Bd. 15, Verl.J.Beltz, Weinheim-Berlin-

Basel 1968 (Herausg.von Ingenkamp, K.-H.u.Marsolek, T.).

- SCHRADER, W.B.: Aufgaben- und Testanalyse als Methode der Testkonstruktion.
Pädagogisches Zentrum, Veröffentlichungen Reihe C:
Berichte, Bd. 15, Verl.J.Beltz, Weinheim-Berlin-
Basel 1968 (Herausg.von Ingenkamp, K.-H.u.Marsolek, T.).
- SCHÜTZ, E.: Studium und Schulleistungen.
DÄ Nr. 9/2.3.1968.
- SCHULZ, U., H.REMSCHMIDT u.H.PRINZ: Leistungstests im Medizin-
studium.
J.F.Lehmanns Verlag, München 1970.
- SCHULZ, U., u.H.PRINZ: Modell eines Informationssystems für
das Ausbildungs- und Prüfungswesen in der Medizin.
Sonderdruck MMW, 111. Jahrg., Nr. 50, 1969.
- STIMPEL, H.M.: In: Prüfungen als Hochschuldidaktisches Problem.
Ergebnisse und Materialien eines Expertenseminars in
Hamburg-Rissen Jan./Febr.1969. In: Blickpunkt
Hochschuldidaktik Heft 1/1969.
- STOLTENBERG, .: Informationsblatt über die Richtlinien der
Med.Fakultäten der Bundesrepublik Deutschland.
Herausgeg.von der Zentralen Registrierstelle für
die Zulassung zum Studium der Medizin und Zahn-
medizin bei der rhein.Friedrich-Wilhelms-Universität,
Bonn, DÄ Nr. 12, 1966.
- TESCHNER, J.: Medizinstudenten auf der Wartebank.
DÄ 48, 1965, S. 2639.
- THURSTONE, L.L.: Current Issues in Factor Analysis.
Psych.Bull.37, 1946.
- ÜBERLA, K.: Faktorenanalyse.
Springer Verlag, Berlin-Heidelberg-New York 1968.
- UEXKÜLL, v.Th.: Probleme der Unterrichtsforschung.
Urban und Schwarzenberg, München-Berlin-Wien 1968.
- VERNON, P.E., u.G.D.MILLICAN: A Further Study of the Reliability
of English Essays.
Brit.J.Statistical Psych. 7, 1954.

WANTMAN, M.J.: Anwendung von Testergebnissen zur Verbesserung des Unterrichts.

Pädagogisches Zentrum, Veröffentlichungen Reihe C: Berichte, Bd. 15, Verl.J.Beltz, Weinheim-Berlin-Basel 1968 (Herausg.von Ingenkamp, K.-H.u.Marsolek, T.)

WISEMAN, S.: Symposium: The Use of Essays in Selection at 11.III. Reliability and Validity.

Brit.J.Educ.Psych.26, 1956.

XI. Anhang

Botanikklausur

1. Äquimolare Lösungen sind isotonisch. Welche Rohrzuckerlösung (Mol.Gew. 342) ist einer 0,9 %igen Traubenzuckerlösung (Mol.Gew.180) isotonisch?
 - A. 1,71 %
 - B. 3,42 %
 - C. 0,171 %
 - D. 0,855 %
 - E. 1,0 %
2. Als antithetischen Generationswechsel bezeichnet man einen Generationswechsel der
 - A. mit einem Wechsel der Gestalt gekoppelt ist
 - B. mit einem Wechsel der Kernphase gekoppelt ist
 - C. über mehr als zwei Generationen läuft
 - D. A. und B. ist richtig
 - E. A. und C. ist richtig
3. Welcher Typus der sexuellen Fortpflanzung findet sich bei Basidiomyceten?
 - A. Anisogamie
 - B. Oogamie
 - C. Isogame Gametangio gamie
 - D. Anisogame Gametangio gamie
 - E. Somatogamie
4. Durch welche Pigmente sind Chromoplasten gefärbt?
 - A. Chlorophylle
 - B. Carotinoide
 - C. Biliproteine
 - D. Anthocyane
 - E. Flavone
5. Welche Substanzen bauen die Stützmembran der Bakterienwand auf?
 - A. Cellulose
 - B. Polygalakturonsäure
 - C. Muropeptide
 - D. Chitin
 - E. Proteine

6. Welche Gärung verläuft aerob?
- A. Alkoholische Gärung
 - B. Milchsäuregärung
 - C. Essigsäuregärung
 - D. Buttersäuregärung
 - E. Keine Antwort ist richtig
7. Welche der nachstehend formelmäßig wiedergegebenen Umsetzungen ist Teilprozeß einer Chemosynthese?
- A. $2 \text{ NH}_3 + 3 \text{ O}_2 \longrightarrow 2 \text{ HNO}_2 + 2 \text{ H}_2\text{O}$
 - B. $\text{C}_2\text{H}_5\text{OH} + \text{O}_2 \longrightarrow \text{CH}_3\text{COOH} + \text{H}_2\text{O}$
 - C. $2 \text{ H}_2 + \text{O}_2 \longrightarrow 2 \text{ H}_2\text{O}$
 - D. A. und B. ist richtig
 - E. A. und C. ist richtig
8. In welchen Strukturelementen der Zelle sind die Enzyme des Citronensäurecyclus lokalisiert?
- A. in den Plastiden
 - B. in den Dictyosomen
 - C. im endoplasmatischen Reticulum
 - D. in den Mitochondrien
 - E. Keine Antwort ist richtig.
9. Aus der Kalorienzahl der Bruttogleichung für die Atmung $\text{C}_6\text{H}_{12}\text{O}_6 + 6 \text{ O}_2 \longrightarrow 6 \text{ CO}_2 + 6 \text{ H}_2\text{O} + 675 \text{ kcal}$ und der Kalorienzahl, die bei der alkoholischen Gärung frei wird $\text{C}_6\text{H}_{12}\text{O}_6 \longrightarrow 2 \text{ CO}_2 + 2 \text{ C}_2\text{H}_5\text{OH} + 21 \text{ kcal}$, ergibt sich die im Äthylalkohol pro Mol enthaltene Wärmemenge zu
- A. 573 kcal
 - B. 654 kcal
 - C. 327 kcal
 - D. 112,5 kcal
 - E. Keine Angabe ist richtig.
10. Saprophyten nennt man solche Organismen,
- A. die mit anderen in Symbiose leben
 - B. die unter völligem Luftabschluß leben
 - C. die auf organischem Substrat leben
 - D. die ihre Energie durch die Oxydation anorganischer Stoffe gewinnen
 - E. die molekularen Stickstoff binden.

11. Die CO_2 -Fixierung in Chloroplasten erfolgt
- A. durch Reduktion über $\text{NADPH}^+ + \text{H}^+$
 - B. durch Anlagerung an einen C_2 -Körper
 - C. durch Reduktion zu Formaldehyd nach der BAEYER-Theorie
 - D. durch Anlagerung an Ribulose-1,5-diphosphat
 - E. auf noch unbekanntem Wege
12. In glykosidischer Bindung liegen in der Pflanzenzelle vor
- A. Amygdalin
 - B. Nicotin
 - C. Digitoxin
 - D. A. und C. ist richtig
 - E. B. und C. ist richtig
13. Durch die Lichtreaktionen und den nicht-cyclischen Elektronentransport wird NADP^+ reduziert,
weil
im Verlaufe des cyclischen Elektronentransportes ATP gebildet wird.
14. Bei Übertragung einer pflanzlichen Zelle in eine hypertoni-
sche Lösung eines Calciumsalzes tritt Konvexplasmolyse
ein,
weil
Calcium-Ionen entquellend auf das Cytoplasma wirken.
15. Bei Eisenmangel zeigen die grünen Pflanzen chlorotische
Erscheinungen,
weil
die Chlorophyllsynthese gestört ist.
16. Die Wasserleitung der Pflanzen erfolgt in Tracheiden und
Tracheen,
weil
die Geleitzellen die erforderliche Energie hierfür liefern.

17. Im turgeszenten Zustand nimmt die Pflanzenzelle kein Wasser mehr auf,
weil
der Wanddruck gleich dem osmotischen Druck ist.
18. Während der Meiose findet eine Neukombination der Gene statt,
weil
der Verlauf der Prophase ein crossing over erfolgt.
19. Der bei der Photosynthese entwickelte Sauerstoff stammt aus dem Kohlendioxyd,
weil
dieses in einem als Photolyse bezeichneten Prozeß durch Licht gespalten wird.
20. Die Sklerenchymfasern gehören zu den Festigungsgeweben,
weil
ihre Zellwände durch in den Kanten verlaufende Verdickungsleisten verstärkt sind.

Statistiken der Botanikklausur

Teilnehmerzahl	:	84
Anzahl der Aufgaben	:	20
Zuverlässigkeit	:	0,4570
Mittelwert der Punktzahlen	:	8,297619
Streuung der Punktzahlen	:	2,683191

Itemstatistiken

Aufgabe Nr.1

Nicht beantwortet	1,19 = 0,00 + 1,19
SI	63,10
TI	0,2044

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	63,10 = 32,14 + 30,95	0,2044
B	5,95 = 5,95 + 0,00	-0,1592
C	16,67 = 11,90 + 4,76	-0,0853
D	13,10 = 7,14 + 5,95	-0,1614
E	0,00 = 0,00 + 0,00	0,0000

Nr.2

Nicht beantwortet	0,00 = 0,00 + 0,00
SI	29,76
TI	0,1881

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	11,90 = 7,14 + 4,76	-0,0682
B	29,76 = 14,29 + 15,48	0,1801
C	8,33 = 3,57 + 4,76	0,0468
D	34,52 = 20,24 + 14,29	0,0594
E	15,48 = 11,90 + 3,57	-0,2805

Nr.3

Nicht beantwortet	0,00 = 0,00 + 0,00
SI	40,48
TI	0,3153

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	3,57 = 2,38 + 1,19	-0,0453
B	11,90 = 8,33 + 3,57	-0,2600
C	23,81 = 17,86 + 5,95	-0,1974
D	20,24 = 10,71 + 9,52	0,0564
E	40,48 = 17,86 + 22,62	0,3152

Nr.4

Nicht beantwortet $0,00 = 0,00 + 0,00$

SI 60,71

TI 0,2618

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
-------------	-----------------------------	----------------------

A	$16,67 = 14,29 + 2,38$	-0,1568
---	------------------------	---------

B	$60,71 = 26,19 + 34,52$	0,2618
---	-------------------------	--------

C	$7,14 = 4,76 + 2,38$	-0,0997
---	----------------------	---------

D	$9,52 = 8,33 + 1,19$	-0,2325
---	----------------------	---------

E	$5,95 = 3,57 + 2,38$	0,1034
---	----------------------	--------

Nr.5

Nicht beantwortet $0,000 = 0,000 + 0,000$

SI 52,38

TI 0,3190

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
-------------	-----------------------------	----------------------

A	$7,14 = 3,57 + 3,57$	-0,0824
---	----------------------	---------

B	$10,71 = 9,52 + 1,19$	-0,1532
---	-----------------------	---------

C	$52,38 = 26,19 + 26,19$	0,3190
---	-------------------------	--------

D	$7,14 = 4,76 + 2,38$	-0,1169
---	----------------------	---------

E	$22,62 = 13,10 + 9,52$	-0,1448
---	------------------------	---------

Nr.6

Nicht beantwortet $1,19 = 0,00 + 1,19$

SI 19,05

TI 0,0818

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
-------------	-----------------------------	----------------------

A	$28,57 = 15,48 + 13,10$	-0,0505
---	-------------------------	---------

B	$11,90 = 7,14 + 4,76$	-0,1093
---	-----------------------	---------

C	$19,05 = 10,71 + 8,33$	0,0818
---	------------------------	--------

D	$4,76 = 3,57 + 1,19$	-0,0040
---	----------------------	---------

E	$34,52 = 20,24 + 14,29$	0,0408
---	-------------------------	--------

Nr.7

Nicht beantwortet $0,00 = 0,00 + 0,00$

SI 17,86

TI 0,3885

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$34,52 = 14,29 + 20,24$	0,2181
B	$17,86 = 15,48 + 2,38$	-0,2718
C	$7,14 = 5,95 + 1,19$	-0,1169
D	$22,62 = 17,86 + 4,76$	-0,2827
E	$17,86 = 3,57 + 14,29$	0,3885

Nr.8

Nicht beantwortet $1,19 = 1,19 + 0,00$

SI 59,52

TI 0,3717

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$10,71 = 7,14 + 3,57$	-0,0815
B	$11,90 = 9,52 + 2,38$	-0,1504
C	$3,57 = 3,57 + 0,00$	-0,0692
D	$59,52 = 26,19 + 33,33$	0,3717
E	$13,10 = 9,52 + 3,57$	-0,2798

Nr.9

Nicht beantwortet $0,000 = 0,00 + 0,000$

SI 17,86

TI 0,2263

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$4,76 = 2,38 + 2,38$	0,1210
B	$48,81 = 27,38 + 21,43$	-0,0639
C	$17,86 = 5,95 + 11,90$	0,2263
D	$9,52 = 5,95 + 3,57$	-0,0662
E	$19,05 = 15,48 + 3,57$	-0,1555

Nr.10

Nicht beantwortet $1,19 = 0,00 + 1,19$

SI 41,67

TI 0,4012

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$19,05 = 11,90 + 7,14$	-0,0764
B	$7,14 = 3,57 + 3,57$	-0,0135
C	$41,67 = 15,48 + 26,19$	0,4012
D	$15,48 = 11,90 + 3,57$	-0,1088
E	$15,48 = 14,29 + 1,19$	-0,3664

Nr.11

Nicht beantwortet $1,19 = 1,19 + 0,00$

SI 30,85

TI 0,4728

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$25,00 = 17,86 + 7,14$	-0,1153
B	$13,10 = 7,14 + 5,95$	0,0227
C	$11,90 = 10,71 + 1,19$	-0,3285
D	$30,95 = 7,14 + 23,81$	0,4728
E	$17,86 = 13,10 + 4,76$	-0,1791

Nr.12

Nicht beantwortet $1,19 = 0,00 + 1,19$

SI 25,00

TI 0,1306

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$35,71 = 21,43 + 14,29$	0,0377
B	$5,95 = 5,95 + 0,00$	-0,2717
C	$15,48 = 8,33 + 7,14$	-0,1088
D	$25,00 = 11,90 + 13,10$	0,1306
E	$16,67 = 9,52 + 7,14$	0,0575

Nr.13

Nicht beantwortet $1,19 = 0,00 + 1,19$

SI 41,67

TI 0,1402

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$20,24 = 11,90 + 8,33$	0,0214
B	$23,81 = 14,29 + 9,52$	-0,1974
C	$41,67 = 21,43 + 20,24$	0,1402
D	$7,14 = 5,95 + 1,19$	0,0381
E	$5,95 = 3,57 + 2,38$	-0,0467

Nr.14

Nicht beantwortet $0,00 = 0,00 + 0,00$

SI 4,76

TI 0,2044

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$41,67 = 29,76 + 11,90$	-0,1927
B	$4,76 = 0,00 + 4,76$	0,2044
C	$21,43 = 8,52 + 11,90$	0,2448

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
D	$29,76 = 15,48 + 14,29$	-0,0431
E	$2,38 = 2,38 + 0,00$	-0,1919

Nr.15

Nicht beantwortet	$1,19 = 0,00 + 1,19$
SI	53,57
TI	0,2990

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$14,29 = 8,33 + 5,95$	-0,0072
B	$8,33 = 7,14 + 1,19$	-0,2903
C	$7,14 = 3,57 + 3,57$	-0,0308
D	$53,57 = 28,57 + 25,00$	0,2990
E	$15,48 = 9,52 + 5,95$	-0,1824

Nr.16

Nicht beantwortet	$0,00 = 0,00 + 0,00$
SI	78,57
TI	0,3282

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$78,57 = 40,48 + 38,10$	0,3282
B	$2,38 = 1,19 + 1,19$	0,0118
C	$10,71 = 8,33 + 2,38$	-0,2966
D	$1,19 = 0,00 + 1,19$	0,0287
E	$7,14 = 7,14 + 0,00$	-0,1858

Nr.17

Nicht beantwortet	$0,00 = 0,00 + 0,00$
SI	82,14
TI	0,3066

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$11,90 = 11,90 + 0,00$	-0,2737
B	$1,19 = 0,00 + 1,19$	0,0287
C	$3,57 = 2,38 + 1,19$	-0,1170
D	$82,14 = 41,67 + 40,48$	0,3066
E	$1,19 = 1,19 + 0,00$	-0,0940

Nr.18

Nicht beantwortet 0,00 = 0,00 + 0,00

SI 27,38

TI 0,3697

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$30,95 = 27,38 + 3,57$	-0,3526
B	$9,52 = 3,57 + 5,95$	0,0094
C	$21,43 = 9,52 + 1,90$	0,0718
D	$27,38 = 8,33 + 19,05$	0,3697
E	$10,71 = 8,33 + 2,38$	-0,1101

Nr.19

Nicht beantwortet 0,00 = 0,00 + 0,00

SI 26,19

TI 0,4587

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$19,05 = 15,48 + 3,57$	-0,2346
B	$32,81 = 14,29 + 9,52$	-0,0308
C	$8,33 = 7,14 + 1,19$	-0,1298
D	$22,62 = 14,29 + 8,33$	-0,1448
E	$26,19 = 5,95 + 20,24$	0,4587

Nr.20

Nicht beantwortet 0,00 = 0,00 + 0,00

SI 57,14

TI 0,4547

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$57,14 = 22,62 + 34,52$	0,4547
B	$1,19 = 1,19 + 0,00$	-0,0531
C	$8,33 = 7,14 + 1,19$	-0,1619
D	$32,14 = 25,00 + 7,14$	-0,3233
E	$1,19 = 1,19 + 0,00$	-0,2167

Zoologieklausur

1. Was bedeutet Chromosomenpolymorphismus?

- A. Die Zahl der Chromosomen nach einer Endomitose.
- B. Die Unterschiede in Länge und Lage des Centromers zwischen Chromosomen eines Karyotyps.
- C. Die Folge einer somatischen Assoziation (Scheinreduktion) bei der Bildung von polytenen Chromosomen.
- D. Verschiedene Chromosomenzahlen innerhalb einer Art als Folge von Translokationen.
- E. Die Schleifenbildung bei der Chromosomenpaarung als Folge von Inversionen.

2. Was sind Pheromone?

- A. Ausscheidungsprodukte am Mantelrand von Mollusken zur Schalenbildung.
- B. Sexuallockstoffe von Eizellen.
- C. Überbegriff für Stoffe, die von Tieren ausgeschieden werden und bei anderen Tieren der gleichen Art bestimmte Reaktionen (z.B. im Sozialverhalten) oder eine bestimmte Entwicklung auslösen.
- D. Enzyme innerhalb eines Lysosoms (oder Nahrungsvakuole) von verdauenden Zellen zum Abbau größerer Nahrungsteilchen.
- E. Genprodukte an polytenen Chromosomen (in den puffs).

3. Ist die Befruchtungsmembran

- A. die primäre Eihülle (= Dottermembran)?
- B. ein Schleimpropf im Uterus, der weitere Besamungen bzw. Begattungen verhindert?
- C. die aus den umgebenden Follikelzellen gebildete Zone, die das Ei bei Wirbeltieren umgibt?
- D. die Masse der inaktivierten Spermien, die durch Fertisilin an der Eihülle festkleben?
- E. keine Aussage ist richtig.

4. Was ist ein Komplexlocus?

- A. Die Stelle eines Chromosoms, an der ein Bruch mit nachfolgender Drehung (Inversion) erfolgt ist?
- B. Die Stelle eines Chromosoms, an der cross-over erfolgt ist?
- C. Die Stelle eines Chromosoms, an der mehrere Pseudoallele benachbart sind?
- D. Eine Schleife am Lampenbürstenchromosom während der RNS-Bildung?
- E. Eine Stelle am Chromosom, die den Nucleolus aufbaut?

5. Wodurch kann hinsichtlich des Chromosomenbestandes ein Mosaikorganismus entstehen, da die Tochterzellen bei einer Zellteilung verschiedene Chromosomensätze zugeteilt bekommen haben?
- A. durch non-disjunction von Geschlechtschromosomen
 - B. durch Inversion an einem Chromosom und nachfolgendem cross-over im Bereich der Inversion
 - C. durch Entstehen eines Isochromosoms nach Querteilung eines Centromers
 - D. durch Entstehen eines Isochromosoms nach vorausgegangener Translokation
 - E. Drei der Möglichkeiten treffen zu.
6. Bei der Kreuzung zweier homozygoter Tiere, die sich in einem Merkmal unterscheiden, ist die Nachkommenschaft der Fl.-Generation nicht einheitlich hinsichtlich dieses Merkmals. Worauf beruht diese Abweichung?
- A. Es handelt sich um ein dominantes Merkmal, für das in einem Teil der Gameten cross-over stattgefunden hat.
 - B. Es handelt sich um ein geschlechtsgebundenes Merkmal, das rezessiv ist, und auf dem X-Chromosom des ♂ lokalisiert ist.
 - C. Es handelt sich um ein geschlechtsgebundenes Merkmal, das rezessiv ist und auf beiden X-Chromosomen des ♂ lokalisiert ist.
 - D. Es handelt sich um die Folgen einer Deletion am Y-Chromosom des ♂.
 - E. Keine der Möglichkeiten trifft zu.
7. Was bedeutet modifikatorische Geschlechtsbestimmung?
- A. Das Tier hat bisexuelle Potenzen und kann je nach Situation als Männchen oder Weibchen funktionieren, ist also zwittrig.
 - B. Es sind zwar Geschlechtschromosomen vorhanden, aber ihre Wirkung wird durch äußere Faktoren unterdrückt.
 - C. Es sind zahlreiche Geschlechtschromosomen vorhanden, und nach ihrer Verteilung in den Gameten entstehen verschiedene Geschlechtstypen.
 - D. Die Bestimmung des Geschlechts hängt von der Einwirkung äußerer Faktoren ab.
 - E. Es ist die Entwicklung abweichender Geschlechtstypen (z.B. Überweibchen) gemeint, die durch Fehlverteilung der Geschlechtschromosomen entstanden sind und die normale Ausprägung der sekundären Geschlechtsmerkmale modifiziert haben.
8. In welchen Tierstämmen lassen sich zum ersten Mal Gruppen von Nervenzellen als pace-maker nachweisen?
- A. Als Riesenfasern (Kolossalfasern) im Bauchmark von Anneliden
 - B. Im Herzen von Fischen
 - C. In den "Kiemenherzen" des Lanzettfischchens
 - D. Im Nervenetz der Ringmuskulatur von Quallen
 - E. Als Riesenfasern des Mantelganglions bei Tintenfischen (zur Koordination der Fluchtbewegung mittels Wasserrückstoß).

9. Wie funktionieren Protonephritiden?

- A. Durch Ansaugen von Gewebsflüssigkeiten durch Zellspalten in den Hohlraum, in dem sich der zugehörige Geißelschopf bewegt.
- B. Durch Ansaugen mittels Wimpern von Coelomflüssigkeit durch einen offenen Trichter.
- C. Durch Weitergabe von geformten Exkretstoffen aus einem geschlossenen Gefäßsystem.
- D. Durch Schleimsekretion, wobei Exkretstoffe in dem Schleim eingelagert werden.
- E. Aussagen C und D sind richtig.

10. Was ist Gamogonie?

- A. Eine spezifische Vielteilungsform bei Sporozoen.
- B. Die Bezeichnung für das Verschmelzen von ♀ und ♂ Vorkernen.
- C. Die Aufteilung einer Protozoenzelle in Gameten.
- D. Das Organ, in dem bei Parenchymiern die Spermien auf die zusammengesetzten Eier treffen.
- E. Der Prozeß, bei dem Urgeschlechtszellen in ein Keimepithel einwandern.

11. Welche Aufgabe hat der Fettkörper bei Insekten?

- A. Bildung der Eiweiße für die Hämolymphe.
- B. Bildung eines Atmungsfarbstoffes für die Hämolymphe.
- C. Speicherung von Exkretionsstoffen und Entgiftung.
- D. Übernahme der Atmungsfunktion, da kein Atmungsfarbstoff vorhanden ist.
- E. Keine Aussage trifft zu.

12. Wodurch werden Polysomen in der Zelle in der typischen Anordnung bei der Eiweißsynthese zusammengehalten?

- A. Durch die Membranen des endoplasmatischen Reticulums.
- B. Durch ribosomale RNS (aus dem Nucleolus).
- C. Durch die entstehenden Polypeptidketten.
- D. Durch die Makromoleküle des Hyaloplasmas ("Cytoskelett").
- E. Durch messenger-Rns.

13. Wie heißen Zellen mit embryonalem Charakter, die bei Schwämmen Zellersatz und Neubildungen bei Regeneration übernehmen?

- A. Gemmulae
- B. Neoblasten
- C. Interstitielle Zellen
- D. Archaeozyten
- E. Collencyten

14. Auf dem 1. Chromosom (= X-Chromosom) von *Drosophila* ist eine Deletion eingetreten, die sich an Speicheldrüsenchromosomen erkennen läßt und dort 30 Querscheiben der Zonen 3 A bis 3 D im heterozygoten Tier an einem der beiden homologen Chromosomen umfaßt. Es wurden also die Loci für "white" und für "notch" erfaßt, aber auch noch zahlreiche benachbarte Genorte. Im Falle einer Kreuzung heterozygoter Tiere treten Zygoten auf, die auf beiden ersten Chromosomen diese Defizienz tragen. Die Embryonen sind letal,
- A. weil eine große Zahl von Genorten fehlt und dieser Verlust nicht ausgeglichen werden kann, so daß die Zygoten absterben
 - B. weil die Allele für "white" fehlen und die Embryonen erst nach 12 Std. Entwicklung absterben
 - C. weil durch die Defizienz die somatische Assoziation der Chromosomen gestört wird
 - D. weil die Allele für "notch" fehlen und die Embryonen nach 6 Std. Entwicklung absterben
 - E. weil durch die Defizienz schon in den ersten Furchungsteilungen die Verteilung der Chromosomen gestört ist.
15. Bei der Betrachtung eines Tieres, das durch seine Segmentierung und die Ausbildung einer starken Cuticula als Arthropode zu erkennen ist, findet sich im vorderen Körperbereich folgende Extremitätenverteilung: Antenne, Mandibel, 1. Maxille, 2. Maxille.
Es handelt sich um:
- A. ein Spinnentier (etwa eine Zecke)
 - B. ein Krebstier
 - C. ein Insekt (etwa ein Käfer)
 - D. einen Wasserfloh (*Daphnia*)
 - E. B und D sind richtig.
16. Bei einem sehr starken Befall einer Schafsherde mit Leberegelern kann es zu einer direkten Übertragung der Parasiten kommen, wenn die Zahl der ursprünglich infizierten Tiere sehr groß war,
weil
die Redien im Darm der infizierten Tiere schlüpfen und mit dem Kot ausgeschieden werden, so daß sie mit der pflanzlichen Nahrung von anderen Tieren aufgenommen werden können.
17. Das Becherauge eines Turbellars nimmt mit seinen Sehzellen Licht wahr,
weil
die aufgenommene Lichtenergie (in Form von Photonen) die Mitochondrien trifft und dort eine erhöhte Atmung bewirkt, so daß verstärkter Aufbau von ATP erfolgt.

18. Die großen Gamonten von Monocystis (Sporozoa) sind haploid,
weil
bei der Bildung der Sporozoite eine Reduktionsteilung stattfindet.
19. Der Nahrungstransport bei den Turbellarien und Trematoden erfolgt durch Wanderzellen,
weil
diese Parenchymier nur über ein geschlossenes Gefäßsystem verfügen.
20. Schwämme sind zur Reorganisation und zum Umbau ihres Nervensystems in jungen Wachstumsphasen in der Lage,
weil
die Collencyten und Archaeocyten das diffuse Nervenetz und die einfacher Sinnesepithelien noch einschmelzen bzw. wieder neu aufbauen können.

Itemstatistik der Zoologieklausur

Zuverlässigkeit der Klausur	: 0,57131
Mittelwert	: 9,976
Streuung	: 3,096
Variationsbreite der SI	: 7,14 - 78,5
Mittelwert der SI	: 49,87
Variationsbreite der TI	: 0,07 - 0,53
Mittelwert der TI	: 0,32

Itemstatistik der einzelnen Aufgaben:

Nr.1

Nicht beantwortet	0,00 = 0,00 + 0,00
Schwierigkeitsindex (SI)	60,71
Trennschärfeindex (TI)	0,3716

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	11,90 = 9,52 + 1,19	-0,1962
B	4,76 = 4,76 + 0,00	-0,1969
C	14,29 = 7,14 + 7,14	-0,1287
D	60,71 = 26,19 + 34,52	0,3716
E	2,38 = 1,19 + 1,19	-0,0492

Nr.2

Nicht beantwortet	1,19 = 1,19 + 0,00
SI	66,67
TI	0,4512

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	10,71 = 9,52 + 1,19	-0,1962
B	4,76 = 4,76 + 0,00	-0,1969
C	66,67 = 21,43 + 45,24	0,4512
D	10,71 = 9,52 + 1,19	-0,2584
E	5,59 = 3,57 + 2,38	-0,1118

Nr.3

Nicht beantwortet	0,00 = 0,00 + 0,00
SI	70,24
TI	0,3986

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$11,90 = 10,71 + 1,19$	-0,2821
B	$8,33 = 7,14 + 1,19$	-0,2036
C	$8,33 = 5,95 + 2,38$	-0,1785
D	$1,19 = 0,00 + 1,19$	0,1426
E	$70,24 = 26,19 + 44,05$	0,3986

Nr.4

Nicht beantwortet	$1,19 = 1,19 + 0,00$
SI	44,05
TI	0,3634

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$5,59 = 5,59 + 0,00$	-0,2734
B	$5,95 = 2,38 + 3,57$	-0,0143
C	$77,38 = 30,95 + 46,43$	0,3634
D	$10,71 = 8,33 + 2,38$	-0,1919
E	$2,38 = 2,38 + 0,00$	-0,0744

Nr.5

Nicht beantwortet	$1,19 = 0,00 + 1,19$
SI	44,05
TI	0,2546

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$27,38 = 16,67 + 10,71$	-0,0987
B	$7,14 = 3,57 + 3,57$	-0,0576
C	$9,52 = 3,57 + 5,95$	0,0156
D	$10,71 = 8,33 + 2,38$	-0,2459
E	$44,05 = 17,86 + 26,19$	0,2546

Nr.6

Nicht beantwortet	$2,38 = 1,19 + 1,19$
SI	26,19
TI	0,3106

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$17,86 = 9,52 + 8,33$	-0,0466
B	$23,83 = 14,29 + 9,52$	-0,1582
C	$26,19 = 7,14 + 19,05$	0,3106
E	$3,57 = 2,38 + 1,19$	0,0222
E	$26,19 = 15,48 + 10,71$	-0,1178

Nr.7

Nicht beantwortet $0,00 = 0,00 + 0,00$

SI 75,00

TI 0,2797

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$9,52 = 2,38 + 7,14$	0,0680
B	$5,95 = 3,57 + 2,38$	-0,1280
C	$2,38 = 2,38 + 0,00$	-0,1753
D	$75,00 = 34,52 + 40,48$	0,2797
E	$7,14 = 7,14 + 0,00$	-0,3263

Nr.8

Nicht beantwortet $0,00 = 0,00 + 0,00$

SI 78,57

TI 0,4082

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$4,76 = 3,57 + 1,19$	-0,1066
B	$0,00 = 0,00 + 0,00$	0,0000
C	$4,76 = 4,76 + 0,00$	-0,1969
D	$78,57 = 32,14 + 46,43$	0,4082
E	$11,90 = 9,52 + 2,38$	-0,3177

Nr.9

Nicht beantwortet $2,38 = 2,38 + 0,00$

SI 38,10

TI 0,5127

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$38,10 = 8,33 + 29,76$	0,5127
B	$27,38 = 21,43 + 5,95$	-0,3832
C	$9,52 = 5,95 + 3,57$	-0,0106
D	$7,14 = 3,57 + 3,57$	-0,0576
E	$15,48 = 8,33 + 7,14$	-0,0711

Nr.10

Nicht beantwortet $0,00 = 0,00 + 0,00$

SI 51,19

TI 0,3540

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$20,24 = 9,52 + 10,71$	0,0326
B	$20,24 = 17,86 + 2,38$	-0,3406

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
C	$51,19 = 15,48 + 35,71$	0,3540
D	$1,19 = 1,19 + 0,00$	-0,1055
E	$7,14 = 5,95 + 1,19$	-0,1621

Nr.11

Nicht beantwortet	$7,14 = 3,57 + 3,57$
SI	7,14
TI	0,0768

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$7,14 = 3,57 + 3,57$	0,0768
B	$10,71 = 5,95 + 4,76$	-0,0346
C	$29,46 = 15,48 + 14,29$	-0,1043
D	$5,95 = 2,38 + 3,57$	0,0507
E	$39,29 = 19,05 + 20,24$	0,0534

Nr.12

Nicht beantwortet	$0,00 = 0,00 + 0,00$
SI	50,00
TI	0,5074

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$3,57 = 1,19 + 2,38$	-0,0607
B	$22,62 = 16,67 + 5,95$	-0,2348
C	$21,43 = 17,86 + 3,57$	-0,3333
D	$2,38 = 1,19 + 1,19$	-0,0492
E	$50,00 = 13,10 + 36,90$	0,5067

Nr.13

Nicht beantwortet	$0,00 = 0,00 + 0,00$
SI	40,81
TI	0,4074

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$3,57 = 3,57 + 0,00$	-0,2264
B	$9,52 = 7,14 + 2,38$	-0,1154
C	$33,33 = 20,24 + 13,10$	-0,1984
D	$48,81 = 14,29 + 34,52$	0,4074
E	$4,76 = 4,76 + 0,00$	-0,1607

Nr.14

Nicht beantwortet $1,19 = 1,19 + 0,00$

SI 41,67

TI 0,2326

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
-------------	-----------------------------	----------------------

A	$28,57 = 15,48 + 13,10$	0,0474
---	-------------------------	--------

B	$4,76 = 1,19 + 3,57$	0,0017
---	----------------------	--------

C	$13,17 = 10,71 + 2,38$	-0,2363
---	------------------------	---------

D	$41,67 = 14,29 + 27,38$	0,2326
---	-------------------------	--------

E	$10,71 = 7,14 + 3,57$	-0,1465
---	-----------------------	---------

Nr.15

Nicht beantwortet $2,38 = 0,00 + 2,38$

SI 23,81

TI 0,0494

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
-------------	-----------------------------	----------------------

A	$8,33 = 5,95 + 2,38$	-0,1785
---	----------------------	---------

B	$22,62 = 13,10 + 9,52$	-0,0969
---	------------------------	---------

C	$23,81 = 11,90 + 11,90$	0,0494
---	-------------------------	--------

D	$11,90 = 4,76 + 7,14$	0,0622
---	-----------------------	--------

E	$30,95 = 14,29 + 16,67$	0,0467
---	-------------------------	--------

Nr.16

Nicht beantwortet $1,19 = 0,00 + 1,19$

SI 27,38

TI 0,2634

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
-------------	-----------------------------	----------------------

A	$13,10 = 8,33 + 4,76$	-0,1338
---	-----------------------	---------

B	$38,10 = 20,24 + 17,86$	-0,0969
---	-------------------------	---------

C	$7,14 = 1,19 + 5,95$	0,1514
---	----------------------	--------

D	$13,10 = 8,33 + 4,76$	-0,2021
---	-----------------------	---------

E	$27,38 = 11,90 + 15,48$	0,2634
---	-------------------------	--------

Nr.17

Nicht beantwortet $1,19 = 1,19 + 0,00$

SI 67,86

TI 0,2746

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
-------------	-----------------------------	----------------------

A	$67,86 = 32,14 + 35,71$	0,2746
---	-------------------------	--------

B	$2,38 = 1,19 + 1,19$	-0,0744
---	----------------------	---------

C	$9,52 = 2,38 + 7,14$	0,0025
---	----------------------	--------

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
D	$17,86 = 13,10 + 4,76$	-0,2574
E	$1,19 = 0,00 + 1,19$	0,0363

Nr.18

Nicht beantwortet $0,00 = 0,00 + 0,000$

SI 46,43

TI 0,3078

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$14,29 = 9,52 + 4,76$	-0,1287
B	$16,67 = 10,71 + 5,95$	-0,1410
C	$16,67 = 8,33 + 8,33$	-0,1100
D	$46,43 = 19,05 + 27,38$	0,3078
E	$5,95 = 2,38 + 3,57$	-0,0631

Nr.19

Nicht beantwortet $1,19 = 0,00 + 1,19$

SI 46,43

TI 0,1536

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$46,43 = 19,05 + 27,38$	0,1536
B	$11,90 = 4,76 + 7,14$	-0,0328
C	$8,33 = 7,14 + 1,19$	-0,1785
D	$14,29 = 8,33 + 5,95$	0,0361
E	$17,86 = 10,71 + 7,14$	-0,1068

Nr.20

Nicht beantwortet $3,57 = 2,38 + 1,19$

SI 50,00

TI 0,5305

Alternative	Häufigkeit der Beantwortung	Leistungskorrelation
A	$17,86 = 10,71 + 7,14$	-0,2072
B	$5,95 = 4,76 + 1,19$	-0,0468
C	$4,76 = 4,76 + 0,00$	-0,1969
D	$17,86 = 11,90 + 5,95$	-0,2875
E	$50,00 = 15,48 + 34,52$	0,5305

"Wechsler-Klausur" vom 29. Juni 1968 (Physik)

Name:.....

In runden Klammern (1) ist die Zahl der erreichbaren Punkte angegeben, in eckigen Klammern $\boxed{1}$ bis $\boxed{2}$ die Zahl der Zusatzpunkte. Bei den Alternativfragen ist die richtige Antwort (evtl. mehrere!) anzukreuzen.

1. Arbeit und Leistung

1.1. Wie ist die Arbeit definiert? (1)

Arbeit = ; A =

Erläuterung : $\boxed{1}$

1.2. Wie ist die Leistung definiert? (1)

Leistung = ----- ; P = -----

Erläuterung : $\boxed{1}$

1.3. Man gebe Arbeitseinheiten an! (1)

1.3.1. 1.3.2.

1.3.3. 1.3.4.

1.4. Man gebe Leistungseinheiten an! (1)

1.4.1. 1.4.2.

1.4.3. 1.4.4.

1.5. Man gebe Energieeinheiten an! (1)

1.5.1. 1.5.2.

1.5.3. 1.5.4.

1.6. Ein Gewichtheber bringt eine Hantel der Masse $m = 90 \text{ kg}$ in $t = 1 \text{ s}$ vom Boden zur "Hochstrecke" ($h = 2,2 \text{ m}$).

1.6.1. Wie groß ist die Arbeit A des Gewichthebers? (2)

A =

1.6.2. Wie groß ist seine Leistung P? (1)

P = -----

1.7. Wie sind der Spezifische und der Molare Brennwert eines Nährstoffes definiert?

1.7.1. Spezifischer Brennwert = ----- (1)

W_B = -----; -----; ----- (1)

1.7.2. Molarer Brennwert = ----- (1)

W_B molar = -----; -----; ----- (1)

2. Kausalität (2)

2.1. Ereignis B geschieht immer nach Ereignis A.

2.2. Ereignis B geschieht manchmal vor, manchmal nach Ereignis A.

2.3. Ereignis A geschieht immer nach Ereignis B.

2.4. Es handelt sich um eine zeitliche Folge.

2.5. Es handelt sich um eine räumliche Folge.

2.6. Im Fall 2.1. ist B Ursache, A Wirkung.

2.7. Im Fall 2.1. ist A Ursache, B Wirkung.

3. Radioaktiver Zerfall

3.1. Was ist Radioaktivität ? (1)

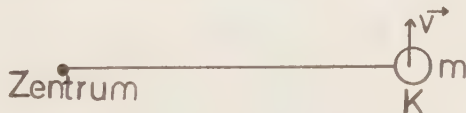
3.2. Nenne einen radioaktiven Stoff! (1)

3.3. Wie hängt die Stärke A eines radioaktiven Präparates von der Zeit ab? (1)

3.4. Ein Präparat der Stärke A ist nach zwei Tagen auf die Hälfte seiner ursprünglichen Stärke abgefallen. Nach welcher Zeit t ist es auf 1/8 der ursprünglichen Stärke abgefallen? (1,5)

4. Kreisbewegung eines Körpers

Der Körper K (Masse m) ist mit einem Faden an einem "Zentrum" befestigt und bewegt sich mit der Bahngeschwindigkeit v auf einem Kreis.



4.1. Man zeichne in obige Figur die Vektoren der Zentrifugal- und der Zentripetalkraft ein und schreibe das betreffende Wort an den entsprechenden Vektorpfeil! (1)

4.2. Wie groß ist die Zentripetalkraft? (2)

F_Z =

Erläuterung der Größen: $\angle I \rceil$

5. Größen und Einheiten

- 5.1. Was ist eine Physikalische Größe? (Was muß man angeben, um die Physikalische Größe eindeutig festzulegen?). (2)
- 5.2. Schreibe die abgekürzten Einheitenzeichen im vollen Wort aus!

(Beispiel: MW = Megawatt = 10^6 Watt). (1)

ns =

mm =

kJ =

Mm =

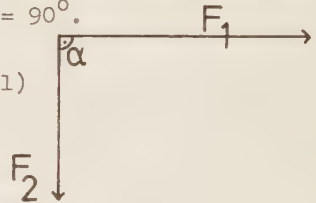
6. Addition von vektoriellen Größen

Zwei Kräfte $F_1 = 6 \text{ N}$ und $F_2 = 8 \text{ N}$ greifen an einem Punkt an; ihre Richtung bilden einen Winkel $\alpha = 90^\circ$.

- 6.1. Man zeichne die resultierende F_R in nebenstehende Figur ein! (1)

- 6.2. Man berechne den Betrag der resultierenden Kraft! (2)

$F_R =$



7. Erhaltungssätze

- 7.1. Nennen Sie einige Erhaltungssätze! (1)
- 7.2. Erläutern Sie an einem dieser Erhaltungssätze dessen Bedeutung! (1)
- 7.3. Ist beim inelastischen Stoß ein Erhaltungssatz verletzt? (2)

8. Viskosität

- 8.1. Geben Sie zwei Methoden zur Bestimmung der Viskosität an! (1
1.) 2.)
- 8.2. Welchen Namen tragen die zugrundeliegenden Gesetze? (1)
- 8.3. Die Stromstärke einer Flüssigkeit bei laminarer Strömung durch ein Rohr der Länge $l = 10 \text{ cm}$ und des Radius $R = 1 \text{ mm}$ hat bei einer Druckdifferenz $p = 100 \text{ Torr}$ den Wert $V = 1 \text{ cm}^3/\text{s}$. Wie groß ist die Stromstärke,
1. wenn der Radius $R = 2 \text{ mm}$ beträgt? (1) $V =$
 2. wenn die Länge $l = 20 \text{ cm}$ beträgt? (1) $V =$
 3. wenn die Druckdifferenz $p = 50 \text{ Torr}$ beträgt? (1) $V =$

9. Aufbau der Materie

9.1. Nennen Sie fünf Elementarteilchen und einige ihrer Eigenschaften! (2,5)

- 1.
- 2.
- 3.
- 4.
- 5.

9.2. Aus welchen Elementarteilchen sind aufgebaut

1. Atomkerne? (1)
2. Atome? (1)

9.3. Geben Sie die ungefähren Größenordnungen an für

1. Atomdurchmesser (1) d_A
2. Kerndurchmesser (1) d_K
3. Abstand der Atome im Kristallgitter (z.B. NaCl) (1) d

10. Ein Eilzug mit der Geschwindigkeit $v = 18 \text{ m/s}$ wird gleichförmig gebremst, so daß er in $t = 15 \text{ s}$ stillsteht.

10.2. Welchen Weg s legt der Zug während der Bremsung noch zurück? (2)

Statistiken der Physikklausur

Zahl der Aufgabenkomplexe : 10
Zahl der Subteste : 50
Maximal erreichbare Punktzahl : 100
MW der erreichbaren Punktzahl : 49
Streuung : 21
Variationsbreite der SI : 29 - 72,3
Variationsbreite der TI : 0,06 - 0,76

Aufgabe Nr. 1

Erreichbare Punktzahl : 14
MW : 5,5
Streuung : 2,7
SI : 39,3
TI : 0,78

Aufgabe Nr. 6

Erreichbare Punktzahl : 3
MW : 2,1
Streuung : 1,1
SI : 72,3
TI : 0,53

Aufgabe Nr. 2

Erreichbare Punktzahl : 2
MW : 1,3
Streuung : 0,66
SI : 66,5
TI : 0,07

Aufgabe Nr. 7

Erreichbare Punktzahl : 3
MW : 2,2
Streuung : 0,97
SI : 61,6
TI : 0,54

Aufgabe Nr. 3

Erreichbare Punktzahl : 4,5
MW : 2,3
Streuung : 1,4
SI : 51,1
TI : 0,29

Aufgabe Nr. 8

Erreichbare Punktzahl : 5
MW : 1,5
Streuung : 0,6
SI : 29
TI : 0,54

Aufgabe Nr. 4

Erreichbare Punktzahl : 4
MW : 2,1
Streuung : 1,3
SI : 61,6
TI : 0,55

Aufgabe Nr. 9

Erreichbare Punktzahl : 7,5
MW : 3,9
Streuung : 1,4
SI : 45,8
TI : 0,55

Aufgabe Nr. 5

Erreichbare Punktzahl : 3
MW : 1,8
Streuung : 0,9
SI : 61,6
TI : 0,48

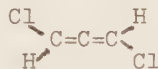
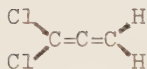
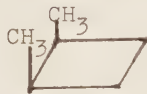
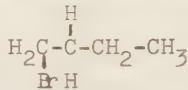
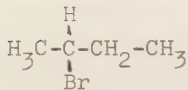
Aufgabe Nr. 10

Erreichbare Punktzahl : 4
MW : 1,9
Streuung : 1,0
SI : 47,5
TI : 0,56

Chemieklausur

1. Formulieren Sie die Reaktionsgleichungen für die Umsetzung von Brom mit n-Butan zu Mono-Brom-Butanen. Wodurch kann die Reaktion in Gang gebracht werden? Welche Isomere werden bevorzugt gebildet und warum?
2. Erläutern Sie an Hand von Beispielen, durch welche Effekte
Carbonium-Ionen
Kohlenstoff-Radikale
Carbanionen
stabilisiert werden können.
3. Formulieren Sie die Reaktionsgleichungen für die Darstellung von Cyclohexen aus
 - a) Cyclohexanol
 - b) Chlor-cyclohexan
 - c) 1,2-Dibrom-cyclohexan.
4. Wie kann man experimentell voneinander unterscheiden
 - a) Hexanal von 2-Hexanon
 - b) 1-Hexen von 1-Hexin
 - c) 1-Hexin von 2-Hexin
 - d) 1-Hexen von 1-Hexanol
 - e) 1-Hexanol von Hexanal
5. Formulieren Sie die Reaktionsgleichungen mit Zwischenstufen für die Addition von HCl an 2-Methyl-2-buten. Welches Isomere wird bevorzugt gebildet?
Begründung.
6. Formulieren Sie die Reaktionsgleichungen für die Bildung von n-Propyl-methyläther.
7. Formulieren Sie die Reaktionsgleichungen für die Darstellung von Äthylbenzol aus Benzol und
 - a) Äthylen (Katalysator : ?)
 - b) Äthylchlorid (Katalysator : ?)

8. Kreuzen Sie von den folgenden Verbindungen die an, welche dissymmetrisch sind, d.h. als Enantiomere auftreten können:



9. Wie kann man primäre, sekundäre und tertiäre Alkohole voneinander unterscheiden?
(2 verschiedene experimentelle Verfahren) Reaktionsgleichungen.

10. Beschreiben Sie die Addition von 5 verschiedenartigen Nucleophilen an Carbonylverbindungen (Aldehyde oder Ketone) durch Reaktionsgleichungen.

Statistiken der Chemieklausur

Anzahl der Aufgaben : 10

Maximal erreichbare Punktzahl: 200 (durch Verdoppelung des
Endresultates)

MW der erreichten Punkte : 69,2

Streuung der erreichten Punkte : 18,8

Variationsbreite der SI : 15,0 - 49,7

Variationsbreite der TI : 0,49 - 0,76

Maximal erreichbare Punktzahl pro Aufgabe : 10

Aufgabe Nr. 1

MW : 4,3
Streuung : 1,0
SI : 43,4
TI : 0,67

Aufgabe Nr. 6

MW : 3,0
Streuung : 0,78
SI : 30,15
TI : 0,5

Aufgabe Nr. 2

MW : 2,6
Streuung : 0,8
SI : 26,15
TI : 0,76

Aufgabe Nr. 7

MW : 3,5
Streuung : 1,0
SI : 35,38
TI : 0,76

Aufgabe Nr. 3

MW : 4,6
Streuung : 0,9
SI : 46,0
TI : 0,64

Aufgabe Nr. 8

MW : 4,96
Streuung : 1,0
SI : 49,7
TI : 0,49

Aufgabe Nr. 4

MW : 2,6
Streuung : 1,1
SI : 26,2
TI : 0,58

Aufgabe Nr. 9

MW : 3,5
Streuung : 0,9
SI : 35,5
TI : 0,62

Aufgabe Nr. 5

MW : 4,1
Streuung : 1,7
SI : 41,3
TI : 0,78

Aufgabe Nr. 10

MW : 1,5
Streuung : 0,5
SI : 15,0
TI : 0,57

Meinem Doktorvater, Herrn Professor Dr.rer.nat.Peter Ihm,
danke ich für die Überlassung des Themas.

Besonderer Dank gilt den Herren Dipl.-Math.Dr.biol.Ulrich Schulz
und Herrn Professor Dr.rer.nat.Dr.med.Helmut Remschmidt für
ihre Anregungen während der Arbeit und für ihre fruchtbare
Kritik bei deren Abfassung.

Lebenslauf

Geboren wurde ich am 24.10.1946 in Paderborn als zweites Kind des Kaufmanns Heinz Broer und seiner Ehefrau Else, geborene Wiethaup. Nach vier Volksschuljahren in Paderborn besuchte ich dort von 1957 - 1966 das neusprachliche Mädchengymnasium St. Michael mit Abiturzeugnis vom 14.2.1966. Mein Studium begann ich an der Philipps-Universität Marburg/L. zunächst mit Psychologie. Nach einem Semester wechselte ich zum WS 1966/67 zur Medizinischen Fakultät über. Die naturwissenschaftliche Vorprüfung bestand ich am 11.9.1967 und die ärztliche Vorprüfung am 28.3.1969. Es folgten 3 klinische Semester in Marburg. Am 24. Sept. 1970 heiratete ich dort den Medizinalassistenten Hartmut Strempel und wechselte Wohnsitz und Hochschule nach Düsseldorf. Hier absolvierte ich 3 weitere klinische Semester und beendete mein Studium am 31.7.1972 mit Bestehen des Staatsexamens. Als Medizinalassistentin kehrte ich nach Marburg zurück, wo ich an der Medizinischen Poliklinik, der Chirurgischen Klinik und der Augenklinik arbeitete.

Meine akademischen Lehrer waren

in Marburg:

Bartenwerfer, Betz, Braasch, Brück, Buchholz, Dimroth, Dölle, Düker, Golenhofen, Graul, Hajos, Hamelmann, Hardewig, Hensel, Heuser, Hort, Karlson, Korb, Linneweh, Lübbers, Mann, Mahr, Martini, Mattes, Milbradt, Merz, Meyer-Bertenrath, Nöth, Nultsch, Petry, Pots, Reim, Rohen, Schleyer, Schmid, Schmitz-Moormann, Seidel (Zool.), Seidel (Chir.), Siegert, Spiegelberg, Straub, Tent, Vogt, Voss, Walcher, Witzel, Zilliken

in Düsseldorf:

Boran, Bay, Beck, Blessing, Büchner, Dolff, Greef, Greiner, Greither, Grün, Haase, Haensch, Hagen, v. Harnak, Hoppe, Kisselbach, Manz, Meiers, Meyer-zum-Gottesberge, Naumann, Pothmann, Schmidt, Strupp, Wirth.

